生命科学名著

〔美〕C.D. Allis
〔德〕T. Jenuwein
〔美〕D. Reinberg
〔美〕M. Caparros
编著

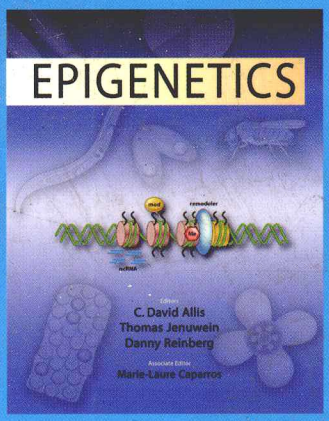
朱 冰 孙方霖 主译

# 表观遗传学

# Epigenetics

科学出版社
www.sciencep.com

(Q-2220.0101)




生命是一个复杂的过程，有很多的例外情况，许多最关键的现象往往不能用经典理论来解释。表观遗传学正是在这些例外的基础之上发展成型的。这些非DNA变化但可继承的现象与癌、衰老、动植物发育等热点问题密切相关，受到广泛关注。

本书是表观遗传学领域领军人物的扛鼎之作，是该领域第一本系统性、权威性论著，不仅涉及一线研究的方方面面，还用专门章节讲述了表观遗传学的发展历程。通过本书，读者可以对这门新兴热点学科有一个深入而完整的认识。

本书适宜于相关领域的研究人员参考之用，作为该领域的教学用书也是非常合适的。

Epigenetics represents a new understanding of biology, besides and beyond the genome. The chromatin template and its function are just as critical as DNA itself for the development of organisms, as well as disease.

——C.D. Allis, T. Jenuwein, D. Reinberg, M. Caparros

生物分社
联系电话：010-64012501
http://www.lifescience.com.cn
e-mail:lifescience@mail.sciencep.com

ISBN 978-7-03-023806-1
9 787030 238061 >

销售分类建议：生物

定 价：98.00 元(含光盘)

生命科学名著

# 表观遗传学

# Epigenetics

〔美〕C. D. Allis
〔德〕T. Jenuwein 编著
〔美〕D. Reinberg
〔美〕M. Caparros
朱 冰 孙方霖 主译

科学出版社
北京

图字：01-2008-2171 号

## 内 容 简 介

生命是一个复杂的过程，有很多的例外情况，许多最关键的现象往往不能用已有理论来解释。表观遗传学正是在这些例外的基础之上发展成型的。这些非 DNA 变化但可继承的现象与癌、衰老、动植物发育等热点问题密切相关，受到广泛关注。

本书是表观遗传学领域领军人物的扛鼎之作，是该领域第一本系统性、权威性论著，不仅涉及一线研究的方方面面，还用专门章节讲述了表观遗传学的发展历程。通过本书，读者可以对这门新兴热点学科有一个深入而完整的认识。

本书适宜于相关领域的研究人员参考之用，作为该领域的教学用书也是非常合适的。

Originally published in English as *Epigenetics* by David C. Allis, Thomas Jenuwein, Danny Reinberg, Marie-Laure Caparros © 2007 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA © 2009 Science Press. Printed in China.
Authorized Simplified Chinese translation of the English edition © 2007 Cold spring Harbor Laboratory Press. Cold Spring Harbor Laboratory Press, the owner of all rights to publish and sell the same.

**图书在版编目(CIP)数据**

表观遗传学/（美）Allis，C. D. 等编著；朱冰，孙方霖主译. —北京：科学出版社，2009

（生命科学名著）

ISBN 978-7-03-023806-1

Ⅰ. 表… Ⅱ. ①A…②朱…③孙… Ⅲ. 发育遗传学 Ⅳ. Q344

中国版本图书馆 CIP 数据核字(2008)第 207034 号

责任编辑：李 晓/责任校对：桂伟利

责任印制：钱玉芬/封面设计：陈 敬

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

http://www.sciencep.com

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2009 年 7 月第 一 版 开本：787×1092 1/16

2009 年 7 月第一次印刷 印张：36 3/4

印数：1—2 500 字数：851 000

**定价：98.00 元（含光盘）**

（如有印装质量问题，我社负责调换〈长虹〉）

# 《表观遗传学》译者名单

**主　译：**朱　冰（北京生命科学研究所）　孙方霖（清华大学）

**翻译和校对人员：**（按姓氏笔画排序）

马　奔　王大亮　王译萱　吕雯雯　朱慧君
刘　岩　杨　鹏　李红艳　吴　慧　张　慧
张道永　郑　泳　赵　磊　赵　璟　袁　文
常建峰　谢精卫

# 译 者 序

对于科学家们来讲，无法被经典理论解释的“意外”现象总是颇具魅力。表观遗传学正是基于对种种“意外”的非孟德尔遗传现象的困惑发展而来。

众所周知，DNA是生命的基本遗传物质，但“令人怦然心动的是，你可以继承的不仅仅只有DNA序列”还有“表观遗传信息”(James Watson 2003)。

在过去的二十多年里，科学家们对DNA和染色质的修饰与重塑、组蛋白变体及RNA等领域不断深入，使得表观遗传学渐成一派，从对若干非孟德尔现象的零星研究汇聚成了一个新兴的学科。表观遗传学与癌症、衰老、干细胞、克隆等研究的密切关系进一步吸引了许多其他学科的科学家和青年学生投入到对该领域的研究，因此，编写一本表观遗传学的教科书也就变得十分必要。2006年，冷泉港实验室出版社邀请活跃于该前沿学科的三位著名学者David Allis、Thomas Jenuwein、Danny Reinberg组织了数十位相关领域的专家，系统编撰了这本涵盖表观遗传学历史、现状和展望的教科书。科学出版社的编辑同志和笔者也希望借此中文版更好地将这门学科介绍给国内的青年学生和其他领域的科学家。

值得指出的是，传统上，我们将“inheritance”及“genetics”均翻译为“遗传”。而表观遗传学研究结果表明，细胞和个体可“继承”的不仅有传统的“遗传”信息，还有“表观遗传”信息。为避免混淆，笔者在本书中将“inheritance”统译为“继承”，这也突出了本学科的宗旨。

笔者及实验室成员承担了本书主体的翻译任务，初稿经审校后再由科学出版社的编辑人员进一步编校成书。我们在此向所有参与本书翻译、校对的人员表示深深的谢意。

由于笔者水平有限，本书难免会有误译之处，欢迎读者批评指正。

朱 冰（北京生命科学研究所）
孙方霖（清华大学）

# 前　言

《表观遗传学》是许多才华横溢的同事们的创作结晶，他们的共同努力使本书得以成稿并让这个编撰过程成为一个富有意义的人生经历。编者们毫不犹豫地向远在伦敦的Marie-Laure Caparros女士致以最高的谢意，没有她的帮助，本书不可能得以最终出版。在本书编撰过程的早期，编者们认识到需要有人协调这一庞大的工作。尤其是需要有专人与40多名作者保持沟通、反馈意见，并让作者们认识到我们需要的不仅仅是他们的专业论述和细致工作。Marie-Laure激励着大家保持创作的动力，勇敢地提出必要的批评意见，提醒着大家保持进度，维持着工作的一致性，使得各章节从雏形发展成终稿。没有她，本书将不可能变为现实。我们也对各自的助手表示感谢，他们是：Elizabeth Conley（David Allis）、Christopher Robinson（Thomas Jenuwein）和Shelli Altman（Danny Reinberg），是他们使我们的工作井井有条，他们是本书的幕后英雄。我们对他们大大小小、不计其数的贡献表示感谢；也感谢他们对我们每个人古怪的工作方式和作为编者的各种缺点所抱有的无尽的耐心。

2004年夏天，在成功的第69届冷泉港“表观遗传学”会议的鸡尾酒会上我们完成了本书雏形的讨论。本书是由冷泉港实验室出版社通过Alex Gann等同事在2003年初委托撰写的，由David Allis、Thomas Jenuwein和Danny Reinberg组织了编者队伍。本书最初的轮廓，包括章节与作者的选择，形成于2003年6月在美国科罗拉多州Snowmass召开的FASEB“染色质与转录”会议的野餐会上。此后我们非常幸运地获得了各领域学术带头专家撰写各个章节的承诺。

在早期策划阶段，我们产生了一个新的思路。我们的每一章节均邀请两位专家撰写，试图为染色质生物学的学生和其他领域的同事提供一整套概述性的介绍，突出那些重要的发现，而非堆砌一些可能很快就会过时的综述。我们希望本书能通过保持一个概述性的框架，而拥有更长久的影响力。此外，通过加入大量的图表和附录，我们希望能尽可能多地列出已知的研究系统和表观遗传标志。每一章前都有一个对本章进行概括的独立成文的摘要，章前也设置了相关图片以吸引读者阅读。

本书的另一重要特点是插图，第3章尤具代表性。Marie-Laure Caparros和Jenuwein实验室（IMP，维也纳）的博士生Stefan Kubicek是创作这些图表的大师。他们对各章节中的图，有些甚至从草图开始，进行了一稿又一稿的润饰，让我们能获得更为一致的表述。来自各个作者实验室的数位博士后和博士生（Gabriella Farkas、Fatima Santos、Heike Wöhrmann等）也为本书贡献了出色的图片。然而，我们并未能将所有图表加以修改。我们还要特别感谢Allis、Jenuwein和Reinberg实验室的Monika Lachner、Mario Richter、Roopsha Sengupta、Patrick Trojer等博士生和博士后对附录表格和摘要的修改、校对和定稿。Steven Gray博士（St. James医院，都柏林）对确认及补充目前所有已知组蛋白修饰的列表给予了极大的帮助。

完成的章节也进行了同行评阅，这些重要的点评使得本书更加流畅并澄清了一些复杂的概念。虽然并非所有的点评都被采纳，但确实对许多章节的形成和全书的整体构架有所帮助。为此，我们对 G. Almouzni、P. Becker、H. Cedar、V. Chandler、W. Dean、R. Feil、A. Ferguson-Smith、M. Gartenberg、S. Grewal、M. Hampsey、E. Heard、R. Metzenberg、V. Pirrotta、F. Santos、T. Schedl、D. Solter、R. Sternglanz、S. Tilghman 等表示深深的感谢。

最后，我们感谢所有作者为本书所贡献的智慧和激情。他们所付出的努力只要看到写就的章节和图表就不言而喻。然而，他们还有着看不见的贡献——他们与编者队伍交流、反馈，引导本书最终成型。概述与概念一章就充分反映了这项贡献，因为在初稿中我们有着太多的个人偏好色彩。我们感谢他们的智慧，使得本书达到更高的境界和平衡。如仍有错误和不足，则完全归因于我们。

本书由冷泉港实验室出版社（纽约）、表观基因组 FP6 NoE 计划（欧盟）、IMP（维也纳）、洛克菲勒大学（纽约）和霍华德-休斯医学院/罗伯特-伍德-强生医学院（Piscataway，新泽西）提供资助。此外，表观遗传学领域试剂和工具的两个领导厂商 Upstate（Lake Placid，纽约）和 AbCam（剑桥，英国）也提供了重要的赞助。

C. D. Allis
T. Jenuwein
D. Reinberg

# 目　录

译者序
前言
第 1 章　表观遗传学：从现象到领域 …… 1
第 2 章　表观遗传学发展简史 …… 17
第 3 章　概述与概念 …… 27
第 4 章　酿酒酵母的表观遗传学 …… 75
第 5 章　果蝇位置效应花斑、异染色质形成和基因沉默 …… 97
第 6 章　表观遗传学研究的真菌模型：裂殖酵母和链孢霉 …… 121
第 7 章　纤毛动物的表观遗传学 …… 151
第 8 章　RNAi 和异染色质组装 …… 179
第 9 章　植物的表观遗传调控 …… 199
第 10 章　染色质修饰及其作用机理 …… 227
第 11 章　Polycomb 蛋白家族调控的转录沉默 …… 251
第 12 章　三胸蛋白家族（trxG）与转录调控 …… 275
第 13 章　组蛋白变体和表观遗传学 …… 295
第 14 章　染色体遗传的表观遗传学调控 …… 315
第 15 章　秀丽隐杆线虫 X 染色体的表观遗传学调控 …… 345
第 16 章　果蝇中的剂量补偿效应 …… 363
第 17 章　哺乳动物中的剂量补偿效应 …… 379
第 18 章　哺乳动物的 DNA 甲基化 …… 405
第 19 章　哺乳动物基因组印记 …… 423
第 20 章　生殖细胞系和多能干细胞 …… 445
第 21 章　淋巴细胞生成的表观遗传学调控 …… 467
第 22 章　核移植和基因组重编程 …… 489
第 23 章　表观遗传学和人类疾病 …… 513
第 24 章　癌症的表观遗传决定因素 …… 539
附录 A　网络资源 …… 561
附录 B　组蛋白修饰 …… 563
附录 C　不同模式生物中的表观遗传机制 …… 570
附录 D　组蛋白修饰的 *Brno* 命名法 …… 571
索引 …… 572

# 第 1 章

# 表观遗传学：从现象到领域

Daniel E. Gottschling

*Fred Hutchinson Cancer Research Center, Seattle, Washington 98109*

## 本章目录

1. 引言
2. 冷泉港会议中的表观遗传学简史
3. 第 69 届会议
   3.1 组蛋白密码假说
   3.2 动态的沉默染色质
   3.3 细胞核的构造
   3.4 朊粒
   3.5 新现象
4. 结束语

致谢

参考文献

## 1. 引　言

2004 年夏，本书的许多作者出席了以“表观遗传学”为主题的第 69 届冷泉港定量生物学会议。作为这次会议的亲历者，我深信这是一次有趣的会议。会议简单地起始于试图对“表观遗传学”作一定义。经过一周的争论，大家意识到这就像是试图给“家庭价值观”下定义一样，每个人都知道它意味着什么，然而每个人又对它有着不同的理解。David Haig 对“表观遗传学”这一词起源的论述也许部分地解释了为何大家对它有着不同的理解。该词在 20 世纪的生物学文献中有两个不同的起源，而其含义至今仍在持续演化。Waddington 在对“基因型决定表型”过程中的“偶然性机制”研究中最先使用了这一名词。此后，Nanney 用这一名词来解释为何具有相同基因型的细胞可以有不同的表型，且这一现象可以持续许多代。

我将表观遗传学现象定义为“非 DNA 突变引起的可继承的表型变化”。而且这一表型变化必须是开关型的，即“开”或“关”，而非程度上的变化。此外，这一改变必须可继承，即便最初引起这一改变的因素已经消失。这样，我认为表观遗传学现象包含以下事件：lambda 噬菌体在裂解型（lysis）和溶原型（lysogeny）间的转变（Ptashne 2004），尿路致病性大肠杆菌的菌毛（pili）形态转变（Hernday et al. 2003），果蝇的位置效应花斑（position-effect variegation）现象（Henikoff 1990），四膜虫可继承的外皮层模式转变（Frankel 1990），朊粒疾病（Wickner et al. 2004a）和 X 染色体失活（Lyon 1993）。

第 69 届会议的召开恰逢冷泉港实验室开展遗传学研究 100 周年，以表观遗传学为主题非常适时。鉴于其历史背景，我想通过对早先冷泉港会议的回顾来解释表观遗传学非常合适。虽然第 69 届会议才第一次以表观遗传学为主题，但表观遗传学研究的内容在这一优秀的会议历史中均有涉及。我将回顾的历史因我个人了解的不周和偏好多有不足。若想要有更完整并严谨地描述，我推荐阅读过去五年来发表的有关表观遗传学的上千篇综述。

通过以下编年史，我希望能表达一种感觉，来阐明这众多的看上去相距甚远的现象是怎样汇聚成一个影响整个生物学研究的领域的。那就是，表观遗传学建立于对解释各种意外现象的尝试，这一点也许比所有其他生物学研究领域都更加鲜明。

## 2. 冷泉港会议中的表观遗传学简史

早在 1941 年第 9 次冷泉港会议上，著名的果蝇遗传学家 H. J. Muller 讲述了他早先报道的对“eversporting displacement”现象的研究进展。先前他用此术语描绘染色体大片段重排后，位于断点附近的基因在个体上出现的嵌合表达现象。在那次会议中，他正式将此现象称为“位置效应花斑”（position effect variegation，PEV）。现在已经很清楚这一现象是因为受影响的基因插入到了“异染色质相邻区域”，从而导致该“常染色质区段”被不同程度地部分“异染色质化”。然而，额外拷贝的异染色质染色体会导

致那些受到影响的基因在功能上趋于正常，这个现象在当时颇为费解，现在我们已经知道这其实是由于它消耗了有限的异染色质组成成分。

1951 年，在第 16 次冷泉港会议上，对基因这一概念的充分理解占据了最重要的位置。所以虽然更多的 PEV 实例被发现，但对此现象的理解却进展寥寥。然而，会议开幕致词点出 PEV 会是未来的一个令人振奋的研究领域（Goldschmidt 1951）。Barbara McClintock 指出染色体位置效应是玉米“可突变基因座”（mutable loci）差异性的原因，并推测她所观察到的玉米突变能力的可变性与果蝇的 PEV 现象可能有着相同的根源（McClintock 1951）。

到第 21 届冷泉港会议时，McClintock 关于其“控制元件”（controlling elements）的观点得到了发展（McClintock 1956）。其中有两点与表观遗传学有着明显的联系。在 *Spm* 控制元件体系中，她发现了一些变种能够让反式因子“抑制”（suppress）某个基因（指减弱或消除其表型），而不是突变该基因。她还指出有些控制元件不但能抑制其插入的基因座，还能抑制与其插入位点两侧有一定距离的基因。其他科学家也讨论了这一“蔓延”现象。J. Schultz 则论述了对异染色质含有量不同的果蝇的生化和生理分析（Schultz 1956）。虽然这一工作非常初步，结论也有限，但它是对异染色质结构进行剖析的早期尝试的代表，也体现了这一任务的艰巨性。

第 23 届冷泉港会议上的两个发言对于我们今天的会议来说是里程碑式的。首先，R. A. Brink 描述了他对玉米 *R* 基因座副突变（paramutation）的惊人发现。$R^{st}$ 和 $R^{r}$ 纯合子有着不同的表型，然而 $R^{st}/R^{r}$ 杂合子自交产生的 $R^{r}$ 纯合子却具有 $R^{st}$ 表型，尽管其并不含有 $R^{st}$ 等位基因（Brink 1958）。但是这一表型却不稳定，会在后几代中回复成 $R^{r}$ 表型。他将这一现象直观地称为副突变，表明这一现象不同于突变，却有相似之处。而后，D. L. Nanney 花了大量的时间来阐述“遗传与表观遗传的概念和作用区别”。简言之，他将表观遗传学赋予了不同于 Waddington 最初提出时的意义。他认为这对于解释其在四膜虫中发现的亲本细胞质能对子代的接合型产生影响是必需的。他的定义同样也适用于其他研究，包括 Brink 在 *R* 基因座的工作和 McClintock 于第 21 届冷泉港会议上提及的工作。

Mary Lyon 早先提出的雌性哺乳动物 X 染色体失活假说（Lyon 1961）在第 29 届冷泉港会议上引起了众多关注。S. Gartler、E. Beutler 和 W. E. Nance 提出了更多的证据来支持这一假说（Beutler 1964；Gartler and Linder 1964；Nance 1964）。Beutler 综述了 X 染色体连锁基因在女性中出现嵌合体表型的多个例证，支持了 X 染色体随机失活的本质。通过对 X 染色体连锁基因产物细致的定量分析，Nance 推断出 X 染色体失活发生于胚胎发育的 32 细胞期之前。

以“染色体结构和功能”为主题的第 38 届冷泉港会议代表了对真核生物染色体研究的回潮。先前由于在原核与噬菌体系统中的研究产生了重大进展，细菌基因表达主导了蓬勃发展的分子生物学。然而，真核生物染色质（DNA 及其组蛋白和非组蛋白）开始得到关注，但是染色质是否在染色体结构、功能还是两者上都具有意义仍不清楚（Swift 1974）。不过，若干研究组开始推测染色质蛋白（包括组蛋白）的翻译后修饰与基因转录或染色体整体结构相关（Allfrey et al. 1974；Louie et al. 1974；Weintraub

1974)。当时关于表观遗传与染色质的相关性只有一丝淡淡的痕迹。由于 McClintock 的控制元件在基因组中重复存在，DNA 重复序列曾经被认为是真核生物多数基因的调控者。然而，有报道表明大多数重复序列与基因无关（Peacock et al. 1974；Rudkin and Tartof 1974)。由于这些报道，重复元件调节基因表达的学说失去了多数与会者的支持。更重要的是，同样是这些研究，发现了大多数重复序列存在于异染色质。

第 42 届冷泉港会议表明，在短短四年中，数量惊人的技术和思想革新彻底改观了对真核生物染色体的研究（Chambon 1978)。这包括限制性内切核酸酶的应用、DNA 重组技术的发展、蛋白质和核酸提取的简化、Southern 和 Northern 杂交的使用、DNA 和 RNA 测序的快速化和染色体免疫荧光技术的使用。核小体的概念被提出、mRNA 剪接被发现，染色质结构，特别是转录活跃区和转录沉默区的生化与细胞学差异，占据了此次会议的中心。然而，与表观遗传最直接相关的是，H. Weintraub 及其同事论述了染色质是如何参与基因表达花斑现象的观点（Weintraub et al. 1978)。

第 45 届冷泉港会议是对 Barbara McClintock 发现“可移动遗传元件”的庆典（Yarmolinsky 1981)。细菌转座子机制研究取得的巨大进展，几乎占据了一半的发言，而其他人则报道了转座或受调控的基因组重排现象不仅在玉米中发生，还存在于各种真核生物中（如果蝇、金鱼草、锥体虫、子囊菌和酵母)。在这样的背景下，所有观察到的基因表达花斑现象都被归因于转座。对“控制元件负责大多数基因调控”观点（Campbell 1981）的认真考问更是出现了沉默，这甚至使某些科学家提出“控制元件的唯一功能是提供遗传可变性”的观点。而异染色质在位置效应花斑现象中的作用受到了质疑。对表观遗传学研究，最值得一提的讨论可能是对酿酒酵母中“沉默交配匣”（silent mating cassette）的深入了解（Haber et al. 1981；Klar et al. 1981；Nasmyth et al. 1981；Rine et al. 1981)。

到第 47 届时，人们对脊椎动物逐渐形成了一个普遍认识，即转录活跃基因的 DNA CpG 甲基化水平低于沉默基因。然而，该现象存在例外，会上也论述的更细致的分析表明启动子区特定区域的甲基化更为重要（Cedar et al. 1983；Doerfler et al. 1983；La Volpe et al. 1983)。基于细菌限制/修饰体系的研究，DNA 甲基化被认为可以阻止重要调节因子的结合。此外，在脊椎动物中 DNA 的甲基化模式能在有丝分裂中被继承。这引出了一个假说，即 DNA 甲基化可能在发育过程中细胞分裂时作为一种转录活性的“记忆”形式（Shapiro and Mohandas 1983)。另一项主要的表观遗传发现是鉴定出了酿酒酵母“沉默交配匣”两侧的 DNA 序列，其存在导致了盒内基因的转录抑制。这是首个被鉴定的参与染色体位置效应的 DNA 序列（Abraham et al. 1983)。

第 50 届会议的主题是“发育的分子生物学”，会上也涵盖了一些重要的进展。被一致认同的最激动人心的进展之一是基本的分子特性在进化上是保守的。例如，人的 RAS 在酿酒酵母中也可以起作用，同源异型盒蛋白在果蝇和人之间是保守的（Rubin 1985)。对染色体印记研究的新尝试开始于小鼠核移植技术的发展（Solter et al. 1985)。这些研究显示，新受精卵的父源和母源基因组中存在有标识其亲本来源性的信息；重要的不仅仅是 DNA，还包括染色体上所含的有关它们亲本来源的额外信息，这种信息为胚胎成功发育所必需。基因的差异性表达依赖于染色体的亲本来源的现象可能揭示了部

分答案（Cattanach and Kirk 1985）。

还有一系列的研究意在弄清双胸复合体（bithorax complex）的调节机制。值得一提的是，E. B. Lewis 特别强调了已知的调节该基因座的反式因子都有令人好奇的共性，即它们几乎全部是该基因座的阻遏因子（Lewis 1985）。因而，在多次细胞分裂过程中维持一个基因的沉默状态，是正常发育所必需的。这与当时的主流想法正相反，即认为基因的活化/诱导是发育过程调控的关键。

DNA 转化和插入突变技术此前在一些物种中取得了进展。其中一项特别有创造性并与表观遗传相关的应用是在果蝇中构建一个含有白眼基因的 *P* 元件（*P* element）转座子并插入整个基因组（Rubin et al. 1985）。这为在果蝇全基因组范围寻找 PEV 发生位点提供了一种方法。

此外，在果蝇（Belote et al. 1985；Maine et al. 1985）和线虫（Hodgkin et al. 1985；Wood et al. 1985）中出现了分析性别决定和性染色体剂量补偿效应的首个遗传学手段也是这次会议的亮点。

第 58 届会议突出了对沃森和克里克发现的 40 周年纪念。纪念的部分内容是众多表观遗传现象的聚会：一些新现象被鉴定，一些现象开始了分子水平分析，一些系统也已取得了足够的进展来提出假说并加以验证。

在锥形虫（trypanosome）中，位于端粒附近的可变表面抗原基因（variable surface antigen gene，*VSG*）家族大都沉默，每一时间段只有一个 *VSG* 表达。尽管这一物种未显示含有甲基化 DNA，有研究报道沉默的 *VSG* 基因含有一种新的稀有碱基：β-D-葡萄糖羟甲基尿嘧啶（Borst et al. 1993）。这一碱基在 DNA 中替代了胸腺嘧啶。很容易得出这与其他物种中胞嘧啶甲基化的相似之处：这两种修饰都对维持基因沉默具重要性。但这一碱基如何被导入 DNA，或者它怎样被赋予这一作用尚不清楚。

对脊椎动物表观遗传现象的研究也取得了进展，这包括染色体印记和 X 染色体失活（Ariel et al. 1993；Li et al. 1993；Tilghman et al. 1993；Willard et al. 1993）。到这次会议时，已清楚地知道哺乳动物中很多染色体位点有印记，在二倍体细胞中只有一个等位基因表达，且表达依赖于其亲本来源。*Igf2-H19* 位点特别有趣，因为它含有两个相邻但调控方式相反的基因。源自父系染色体的 *Igf2* 表达而母源的被抑制，*H19* 则是父源的被抑制而母源的表达。更有趣的是，在父系染色体这两个基因的上游观察到可甲基化的 CpG。有假说认为差异性甲基化调控着两个基因与相邻增强子的结合。该增强子更接近 *H19*，且紧邻于其下游（Tilghman et al. 1993）。因此设想这两个基因排他性地竞争与增强子的结合；当 *H19* 被甲基化时，自由增强子可以激活相对更远距的 *Igf2* 基因。DNA 甲基化在此过程中起调控角色的证据来自于小鼠的研究。在 ES 细胞中将首个编码 5-甲基胞嘧啶 DNA 甲基转移酶的基因突变显示，在胚胎发育过程中，父源的 *H19* 低甲基化且该基因活跃转录（Li et al. 1993）。

对$^{5Me}$CpG 如何起调节作用的一个重要发现是第一个$^{5Me}$CpG DNA 结合复合体（MeCP1）被纯化出来（Bird 1993）。MeCP1 不仅结合 DNA，而且当与一个报告基因上游序列结合时，MeCP1 导致该基因被阻遏。尽管这并不能解释 *Igf2-H19* 位点的调控，但它确实提供了一个潜在机制来说明 DNA 甲基化与基因阻遏之间的普遍相关性。

多年来，遗传定位研究发现人类 X 染色体的一段对 X 染色体失活具关键作用。对 X 染色体失活中心的分子克隆研究导致 *Xist* 基因的发现（Willard et al. 1993），一段约 17 kb 的非编码 RNA 只在失活的 X 染色体表达。小鼠的 *Xist* 基因与人的在结构和序列上具有令人惊奇的同源性，这提供了一个优秀的模式系统，用以剖析该 RNA 通过何种途径实现对大部分 X 染色体的抑制。

链孢霉中有两项值得注意的发现（Selker et al. 1993）。首先，胞嘧啶 DNA 甲基化的发生并不局限于 CpG 双核苷酸处，它似乎发生在任何 DNA 序列中。第二项发现则是对重复序列诱导的点突变（Repeat induced point mutation，RIP）现象令人惊奇的描述。当在单倍体基因组中发生 DNA 序列倍增时（连锁的或非连锁的），该序列就会发生重复序列诱导的突变，并可经接合作用通过生殖周期传递到子代。在此过程中有两个事件发生：一是倍增 DNA 序列两个拷贝同时发生 G∶C→A∶T 的突变，二是重复序列诱导的点突变（RIP）的序列附近的数百个碱基对内发生 DNA 甲基化。这种对基因组的双重攻击极为高效——约 50%的非连锁位点发生重复序列诱导的点突变（RIP），而紧密连锁位点的发生概率近 100%——从而轻易地破坏其基因功能。

当果蝇中的 *brown* 基因易位至异染色质附近时，表现出显性 PEV；易位的拷贝能抑制野生型拷贝的表达。在寻找这一反式失活现象的增强因子和抑制因子的过程中，Henikoff 发现，位于异染色质附近的基因倍增增加了对正常基因拷贝的抑制水平（Martin-Morris et al. 1993）。尽管这一现象的机制还不清楚，人们猜想该现象可能与链孢霉的 RIP 相似，但前者在没有 DNA 甲基化的情形下发生，因为甲基化不存在于果蝇中。

Paul Schedl 阐述了染色体“边界元件”（boundary element）的概念（Vazquez et al. 1993）。第一个被发现的边界元件位于果蝇一热激位点的蓬松（puff）区域两侧，其发现归因于其不寻常的染色质结构：一段约 300bp 核酸酶不敏感的核心区，而其两侧均为核酸酶高敏感位点。这些元件被推测能在染色体上划分染色质区域。有两项体内实验支持该假说：①当边界元件位于报告基因两侧时，这些元件能有效消除报告基因随机插入基因组时常见的染色体位置效应；②边界元件的界定也含有其对增强子的抑制作用。当边界元件被插至某基因的启动子和增强子之间时，该基因的表达被抑制。在其他物种中，边界元件的概念虽然还未被清晰地定义，但也在发展，尤其是哺乳动物的珠蛋白位点（Clark et al. 1993）。

酿酒酵母为研究染色质相关的表观遗传学现象的发生机制提供了良好的系统。沉默交配位点的沉默子被证明是一些 DNA 结合蛋白的识别位点。其结合的效应表现出环境依赖性，如 Rap1 蛋白不仅对基因沉默有重要作用，也结合于一系列基因的上游以激活转录（综述见 Laurenson and Rine 1992）。

多年来，在 DNA 复制和基因组的转录静止区域之间建立了很多联系。失活的 X 染色体、异染色质和沉默的印记位点在 S 期的复制均晚于基因组中的转录活跃区。此外，沉默交配位点的基因沉默的建立需要经过 S 期，表明沉默的染色质必须由新复制的 DNA 来构建。这就使得以下发现变得非常有趣——沉默子是 DNA 复制的起始位点，且其起始活性与沉默功能不可分割（Fox et al. 1993）。并且，在近期被发现的复制起

始识别复合物（origin recognition complex，ORC）的突变体中，它表现出沉默缺陷（Bell et al. 1993；Fox et al. 1993）。

研究发现酿酒酵母端粒与果蝇中的相似，表现出 PEV，为仔细研究异染色质结构及其对基因表达的影响提供了另一视角。插至端粒附近的报告基因导致同一菌落的表达混杂性。阻遏状态所依赖的很多基因（*SIR2*、*SIR3*、*SIR4*）与沉默交配位点基因沉默所必需的基因相一致。对沉默染色质结构以及花斑效应的调控的一些重要方面已经被阐述。值得一提的是，异染色质在细胞学上被定义为致密化的染色质，但酿酒酵母的沉默染色质从未观察到此现象，不过，因其有与果蝇相似的 PEV 现象，我们倾向于认为酵母的沉默染色质是异染色质的等价物（Weintraub 1993）。

通过对酵母的研究，一系列基本概念开始成形。首先，组蛋白 H3 和 H4 的重要性变得清晰。特别是组蛋白 H3 和 H4 的 N 端尾巴与沉默的异染色质的形成直接相关（Thompson et al. 1993）。这些组蛋白尾端的特定突变体能消除沉默或使沉默不完全，从而引出这样一种看法，即认为尾部残基的净电荷与尾部的某些特定残基均作用于基因沉默。另外，早期的染色质免疫沉淀（ChIP）显示，位于组蛋白 H4 氨基末端尾部的赖氨酸，与基因组其他部分相比，在沉默染色质区域表现为低乙酰化。而且，其中一个组蛋白突变体说明 H4 K16（能被乙酰化），对沉默染色质的形成有关键作用。

端粒为研究 Sir 蛋白调控基因沉默提供了一个最简单的系统。招募沉默蛋白（recruiting silencing protein）的概念得以发展。简言之，通过酵母双杂交方法，发现端粒结合蛋白 Rap1 与 Sir3 和 Sir4 有相互作用（Palladino et al. 1993）。因而，Rap1 能“招募”这些 Sir 蛋白到基因组的端粒区。有证据显示，Sir3 和 Sir4 能相互结合，而且最重要的是，Sir3（可能还有 Sir4）与组蛋白 H3 和 H4 的尾端有相互作用（Thompson et al. 1993）。此外，过表达 Sir3 会导致其从端粒沿染色质纤维向内“蔓延”，说明 Sir3 是沉默染色质的限制性成分并能沿染色质发生“多聚化”（Renauld et al. 1993）。总而言之，可能存在一个对于基因沉默具有重要意义的大型相互作用网络：由于其与蛋白 Rap1 的相互作用，Sir 蛋白开始在端粒 DNA 处的装配，而后可能借助与组蛋白 H3 和 H4 尾端的结合，从端粒处沿染色质纤维多聚化。

端粒表达花斑性中的转录状态的转换，可能是基因表达沉默和活跃间竞争的结果（Aparicio and Gottschling 1994；Weintraub 1993）。如果一个端粒位置的基因的转录激活因子缺失，该基因的本底转录机制就不足以表达而使该基因发生组成型沉默。反之，转录激活因子的过表达导致端粒位置的基因组成型表达，即该基因永不沉默。在缺失 *SIR3*（或 *SIR2* 或 *SIR4*）的情况下，本底基因表达足够高，但若增加 *SIR3* 的量，则会相应增加沉默细胞的比例。尽管通过细胞周期，转录激活因子能压制基因沉默，当细胞处于 S 期尤为有效，该时期正是染色质被复制的阶段，因此对竞争最为敏感。有些出人意料的是，处于 $G_2$/M 期的细胞也很容易转换，表明沉默染色质在此时尚未完全被装配。

酵母的沉默染色质表现出对核酸酶和 DNA 修饰酶的抗性，说明与基因组其他部分相比，其所含 DNA 更不易被接近（Thompson et al. 1993）。

看起来在酵母基因组中存在着不同等级的沉默：端粒最易被干扰，*HML* 次之，

*HMR* 则最不敏感。事实上，当 *SIR1* 基因被突变后，正常情况下完全沉默的 *HM* 位点表现出表达的花斑性（Pillus and Rine 1989）。

最后 Sir3 和 Sir4 定位于细胞核边缘区域，而端粒也是在此。有假设认为细胞核的构成使得核边缘为沉默提供了一个特殊环境（Palladino et al. 1993）。

裂殖酵母也含有沉默交配匣，据推测与酿酒酵母沉默交配匣的效果类似。不过，裂殖酵母接合型转变的故事多了一点波折。在一系列精巧实验中，Amar Klar 提出了在细胞中某个“标记”是如何在 DNA 的一条链上留下印记的构想（Klar and Bonaduce 1993）。当双链断裂促使接合型发生转变时，经过两次细胞分裂，该标记存在于四个孙细胞中的一个。裂殖酵母并没有任何已知的 DNA 修饰（甲基化等），因此推测是一种不同类型的印记留存于 DNA 链上。

第 59 次大会的主题是“肿瘤分子遗传学”。肿瘤抑制基因的构想成立以后，肿瘤形成中表观遗传调控的概念得以发展。已有一些研究支持该观念，但一有趣的变化发生于研究 Beckwith-Wiedemann 症和 Wilms 肿瘤患者的过程中。两种患者的突变被定位于一含有印记的 *H19-IGF2* 基因的位点。Feinberg 等在受影响的患者中发现了这些基因的“印记丢失”（loss of imprinting）现象：母源位点丢失其印记，*H19* 被抑制而 *IGF2* 得以表达。因此，原则上在基因组其他部位也可以发生 LOI 现象，可能导致肿瘤形成中的重要基因出现双等位基因同时表达和（或）表达抑制。

到主题为“转录机制”的 63 次大会前的数年间，对一些表观遗传现象分子层面的理解取得了一些重要进展。发现了组蛋白的修饰酶，准确说是组蛋白乙酰化酶和去乙酰化酶。其中一些酶在基因表达调控中扮演重要角色。这为研究直接影响 PEV 和沉默的基因产物提供了一条途径。这些发现在这次大会上露出了冰山一角（Losick 1998）。对酵母中沉默蛋白 Sir3 和 Sir4 的精细研究显示了它们之间相互作用的多价性，也揭示了所有 Sir 蛋白、组蛋白以及各种 DNA 结合因子之间的相互作用网络是如何装配沉默染色质的。此外，各种不同位点（端粒、rDNA、*HM* 位点、双链断裂点）怎样竞争有限的 Sir 蛋白的细节也得以揭示。削弱某一特定位点招募沉默因子的能力，在其他位点的 Sir 蛋白水平得到提高（Cockell et al. 1998）。这为某一位点的沉默会影响其他位点的表观遗传沉默且受浓度作用的影响的观点提供了直接证据，该观点最初来源于对果蝇 PEV 的研究，此前一直未得到验证（Locke et al. 1988）。

另一项发现则解释了 DNA 甲基化是怎样借助染色质来调控基因表达的。这来自对 MeCP2 蛋白复合物的鉴别，MeCP2 能结合甲基化的 DNA 和组蛋白去乙酰化酶（Wade et al. 1998）。甲基化的 DNA 可能招募去乙酰化酶至某位点，从而促进相邻基因的沉默。

边界元件的概念被从果蝇延伸到哺乳动物。来自 β 珠蛋白位点的直接证据表明，在后生动物（也可能在所有真核生物）中，染色质边界元件确实都保守存在（Bell et al. 1998）。

主题为“免疫系统中信号传导和基因表达”的第 64 次会议提供了单等位基因表达如何发生，及其发生可能比原先想象的更为广泛分布的证据。淋巴细胞中免疫球蛋白位点的单等位基因表达已经被观察到了相当一段时间，该现象保证每一个淋巴细胞只产生

一种受体类型（Mostoslavsky et al. 1999）。等位基因的表达在发育早期被随机选择：两个等位基因开始处于抑制状态，但一段时间后，其中之一被去甲基化。虽然还不清楚基因是如何被选择成为单等位表达基因，但在单等位基因表达的必要性并不明显的其他位点也有此现象。如编码白细胞介素 IL-2 和 IL-4 的基因只有一个等位基因表达（Pannetier et al. 1999）。

第 65 届大会上与表观遗传学相关的最重要演讲是发现 Sir2 蛋白是一种组蛋白去乙酰化酶（Imai et al. 2000）。这是唯一在所有真核生物中都具有明确的同源蛋白并调节 PEV 的 Sir 蛋白。它似乎是去除沉默染色质上组蛋白乙酰基的主力军。而且由于它的 NAD 依赖性，将细胞生理状态与沉默（异染色质）调节相联系起来。

议题为“人类基因组”的第 68 届会议是遗传学领域的一个重要里程碑。虽然还有许多遗传学工作要做，但人类基因组和其他物种基因组测序的完成，意味着是进入“超遗传学”（即表观遗传学的字面含义）的时候了。

这一历史性回顾突出了与其他许多研究领域共享的几个主题。首先，它表明了表观遗传学进展的不连续性。其次，随着表观遗传现象的分子机理正在开始被了解，表观遗传学与普遍的生物调控更容易被联系起来。最后，它表明那些被我们认为是科学先贤的人早已指出了这个联系的存在，只是对大多数其他人，需要花更多时间才能看到这些明显的联系罢了。

## 3. 第 69 届会议

多年来，一些通用于所有表观遗传现象的普遍概念得以确定，并对深入探索的实验思路起着指导作用。首先，两种表型状态（“关闭”和“开启”），总是能在其各关键调控点找到结构差异，即形式决定功能。因而，解析这两种不同的结构和其组成成分并找到它们组成上的差异是首要的任务。其次，在充斥着竞争因素和多种力量的环境下，这种不同的结构必须具备维持和延续的能力。因此，要求每种结构具有自我增强能力或正反馈通路，以保证其经过多次细胞分裂仍能被维持和传播；在某些情形下能保持一生，如 X 染色体失活。

许多以前会议中被定义的机制性原理，在第 69 届会议上继续得到改进，同时也有新进展。要在文中阐述这些新进展，很重要的一点得指出两项对表观遗传学具有较大影响的发现。一是 RNA 干扰和相关的以 RNA 为基础的调控机制的发现。另一个则是朊粒假说机制的阐明。上述两个领域在过去十年间进展迅速，其中一些研究有助于了解以染色质为基础的表观遗传学，另一些则为理解表型的可遗传性提供了新视角。

这次会议上报告的许多成果在本书各章中有详细描述，因此我就不在这儿讨论了。然而，我会触及几个我喜欢但未在本书其他章节被提及的进展。最后，我会试图概括我从这次会议得到的最重要的概念。

### 3.1　组蛋白密码假说

考虑到组蛋白修饰及其潜在的信息内容，存在着许多关于“组蛋白密码假说”的讨

论（Jenuwein and Allis 2001）。该假说的支持者引用了如组蛋白 H3 在 K9 位点的三甲基化和其与 HP1 家族异染色质蛋白质极强的亲和力的例子（Jenuwein and Allis 2001）。反对方则引用生化和遗传的证据，如组蛋白 H4 氨基末端尾巴的净电荷，不管该电荷位于何处，对 DNA 结合或表型都有显著的影响（Megee et al. 1995；Zheng and Hayes 2003）。

在酿酒酵母中，运用专一性抗体和染色质免疫沉淀结合基因芯片分析（ChIP-Chip），Grunstein 叙述了全基因组范围组蛋白乙酰化修饰和染色质结合蛋白的分析数据（Millar et al. 2004）。其研究聚焦于同 H4K16 乙酰化相关联的表观遗传转变，导致结合或不结合特定染色质蛋白，从而支持了组蛋白密码假说。尽管未经讨论，他的一些数据似乎也支持了他人报告的结论，即对基因组的大部分而言，特定组蛋白修饰和基因表达之间没有相关性（例如，所有活性基因具有同样的标记，而非活性基因却没有）（Schubeler et al. 2004；Dion et al. 2005）。将所有结果汇总起来，我猜测特定修饰作用和总体净电荷作用，都作用于对染色质结构和基因表达调控的机制中。

## 3.2 动态的沉默染色质

我必须承认，基于异染色质的静态图像和沉默染色质的稳定特性，我曾经确信异染色质状态一旦被建立，就坚如磐石。只有到 DNA 复制的时候，这种不受影响的稳定结构才变得松弛。由于这样的思维方式，我愚蠢地忽略了大学化学所学的动态平衡原理。不过，因为对沉默染色质和异染色质的研究，这些课程又被带回家重温。研究显示，哺乳动物细胞中的异染色质蛋白（HP1）、酵母中的沉默蛋白（Sir3）均处于动态平衡状态，即在异染色质状态和可溶状态之间进行迅速转换，即便异染色质处于最不易受影响的状态时也是如此（Cheng and Gartenberg 2000；Cheutin et al. 2003）。对这种动态特性的认识，迫使人们从不同的角度来思考表观遗传染色质状态是怎样被保持和遗传的。这暗示，在一些系统中，表观遗传状态能在任一时刻被逆转，而不仅仅是在 DNA 复制阶段。因此，我们可以推断沉默染色质自我增强和传播的机理是持续作用的。

组蛋白甲基化被广泛认为是一种能真正给予染色质“永久”标记的修饰（Kubicek and Jenuwein 2004）。与其他所有组蛋白修饰（如磷酸化、乙酰化、泛素化）相反，尚未有已知的酶能够可逆地将一个甲基从赖氨酸或精氨酸的胺基处移去*。而且，在生理条件下借助简单的水解作用去掉甲基基团在热力学上是不利的，因而不大可能自发发生。

那些认为甲基化标记是永久性的观点目前已被数个报道所动摇。研究显示，一种细胞核肽基精氨酸瓜氨酸化酶（使精氨酸转化为瓜氨酸的酶）（PAD4）能从组蛋白 H3 的精氨酸残基上去除单甲基基团（Cuthbert et al. 2004；Wang et al. 2004）。尽管这一去甲基化过程导致精氨酸残基转变成瓜氨酸，并非真正意义上的去甲基化，但为永久性甲基化标记的去除提供了一种机制。

Robin Allshire 提出了一个吸引人的论点，即裂殖酵母的 *tis2* 基因能逆转 H3K9 上

* 目前已经有多种组蛋白去甲基化酶被发现。——译者

的双甲基化（R. Allshire，个人通信）。他或许走在正确的道路上，因为会议之后几个月，发现了一个源自哺乳动物的无关的 LSD1 酶，能特异地使双甲基化或单甲基化 H3K4 去甲基（Shi et al. 2004），从而逆转一“活跃”的染色质标记。有趣的是，LSD1 对三甲基化的 H3K4 不起作用，因此，甲基化能在标记过程中被逆转，但一旦标记完全成熟，则不能逆转*。

不过，Steve Henikoff 报道了一种能去除赖氨酸永久性三甲基化的方法。他显示了以一种独立于复制的转录依赖性的方式，组蛋白变体 H3.3 能取代常规 H3（Henikoff et al. 2004）。大体上，用于沉默的包含甲基化标记的组蛋白能被更有利于转录的另一个组蛋白所置换。当总染色质被分离时，组蛋白 H3.3 较常规 H3 含有更多活跃染色质的甲基化标记（如 K79me）。

所有这些结果表明，可能并不存在一种用于标记的简单的组蛋白修饰，使沉默染色质状态在细胞分裂中得以遗传。而是，肯定存在着一系列更精细的相互作用，它们虽然并不能确保沉默，但能提高保持沉默状态的可能性。

## 3.3　细胞核的构造

许多年前就已发现基因在细胞核内所处位置与表达间有相关性（Mirkovitch et al. 1987）。这个发现导致一种观点——在细胞中存在着特殊的区室，在这些区室中基因的表达或沉默受到限制。有人认为这种组织方式的必要意义是为了保持基因组的复杂性并以一种可行的规则来对其进行调控。酿酒酵母的研究支持该观点，作为沉默复合物关键组分的端粒常位于细胞核周边，Sir4 也是如此（Palladino et al. 1993）。使端粒或 Sir4 离开核周边会导致端粒沉默的缺失（Laroche et al. 1998；Andrulis et al. 2002）。而且，人为地将一个部分沉默的基因引至核周边会使其变得完全沉默（Andrulis et al. 1998）。

在一个重要意义的实验中，Gasser 发现如果将端粒和沉默复合物都从核周边释放，并自由移动于细胞核，端粒沉默仍能有效建立（Gasser et al. 2004）。这表明，没有特别的需要将基因限定于某一区域。这也和下述发现更为一致，即尽管染色质蛋白与染色质快速结合和分离，但仍能有效地进行调控（如沉默）。或许某些定位对特定条件（胁迫?）下保持相关因子的局部高浓度是必需的，也可能是漫长的进化过程产生区域无效组合的一个败笔。

## 3.4　朊粒

Wickner 提出了朊粒的概论和界定标准，其描述清楚地表明朊粒是一种表观遗传现象（Wickner et al. 2004a，b）。从最简单的分子意义而言，朊粒通过改变其同源基因产物，从而导致可遗传的表型改变。朊粒引起其底物的结构改变，这并无 DNA 序列变化发生；有一类被研究得最为透彻的朊粒可使一可溶蛋白变成淀粉状纤维。在许多情形下，淀粉状纤维形式会降低或消除蛋白质的正常活性，从而产生表型的变化。Wickner 定义了另一类不会形成淀粉状的朊粒，它们是一类需要自我激活的酶。如果一个细胞只

* 目前已有多种作用于三甲基化赖氨酸的去甲基化酶被发现。——译者

含有这种酶的非活性形式，那么要求外源的活性形式的酶来启动这个“自我繁殖”的过程，只要细胞中含有一个活性分子，这一现象就能得以持续。Wickner 提供了两个例子，并预期这类蛋白会界定一个新的表观遗传机制以供研究。

Si 提出了初步的证据说明一种朊粒模型可能可以解释海蜗牛的习得记忆（Si et al. 2004）。神经元细胞中一系列储存的 mRNA 转录物，对于其短期记忆的维持很重要。他发现一个蛋白质翻译的调控因子 CPEB 能以两种形式存在，其中 CPEB 的活性形式对其自我增殖起决定性作用。对这一设想的检测尚处于早期阶段，但这为思考我们是如何记忆的问题提供了一个令人兴奋的新途径。

### 3.5 新现象

对于新的、非预期的现象的描述，总是占据着我们的想象力。会后数个星期，我的头脑尤其被一个报道所占据。对 *HOTHEAD* 基因（在拟南芥中调节器官融合）突变等位基因的经典遗传分析显示，其并不遵循正常的孟德尔遗传定律（Lolle et al. 2005）。研究显示，杂合子植株 *HOTHEAD/hothead* 自交产生 *hothead/hothead* 纯合子植株，而 *hothead/hothead* 纯合子植株进行自交，后代中有 15%的 *HOTHEAD/hothead* 基因型。这一惊人水平的野生型回复，在核苷酸水平准确地产生了存在于早几代中的野生型基因的拷贝。这种逆转并不局限于 *HOTHEAD* 位点，其他位点也有相同频率的野生型回复发生。不过所有这些回复均要求父母为 *hothead/hothead* 纯合子。*HOTHEAD* 基因的产物并未对该现象的产生原因提供显而易见的解释，但相关讨论暗示野生型基因的拷贝，也许通过 RNA 传至后代。虽然可能有争议认为因为有 DNA 序列的改变，该现象不属于表观遗传学，但该假定的旧有拷贝的遗传传递并不遵循通常的遗传定律。无论如何，该现象对遗传学有巨大意义，特别是从进化的角度来看。

## 4. 结 束 语

要解释各表观遗传机制还需要做些什么？多数情况下，我们仍然在收集（发现）其组分。就像全基因组测序极大地促进了遗传学的发展，当所有组分都已知时，对表观遗传学很可能会有更为清晰的了解。令人鼓舞的是，过去十年我们已经取得了巨大的进展。

我承认我不能确定我们是接近还是远未达到，对表观遗传状态怎样保持和传代的机理有一个准确的了解。基于朊粒的现象也许会是最先被认知的，而那些基于染色质的现象则可能最后。在染色体上建立沉默状态必需的多价相互作用，增加了问题的复杂性。而沉默染色质的动态性质则使之更为复杂。要想进一步了解各组分进出染色质结构的运动情况，需要应用改良的或新的方法。尽管染色质免疫共沉淀对于确立组分是否存在于某一结构中具有重要意义，但它曾一度蒙蔽了人们关于动态本质的了解。

考虑到复杂性，我猜测，简单地检测所有组合物之间的结合与平衡常数，以及试图导出一系列调控表观遗传转换的微分方程，可能并非是对资源的有效利用，也并不一定导致更好的理解。更确切地说，为了最终理解表观遗传现象，我推测需要发展出一种新

型的数学方法，并与新的实验检测方法相结合。其中部分可能需要完善体外系统，使之能忠实再现不同表观遗传状态间的转换。

大部分表观遗传现象中两种状态之间的较量，可能反映着细胞内多个水平竞争的结果，以及试图修正“间接损害”的尝试。例如，沉默蛋白质可能参与保护基因组免受转座子的影响。但由于沉默蛋白质通过普遍存在的核小体起作用，一些关键性的基因受到抑制。为了克服这一矛盾，组蛋白修饰作用（如 H3K4 和 H3K79 的甲基化）和各种组蛋白变体（H2A. Z）参与阻止沉默蛋白质与关键性基因的结合。从后续的结果来看，这些变化可能介入其他过程，如沉默蛋白质对某些基因的抑制可能是有用的（沉默交配匣）。沉默机制也可能有其他功能，如促进染色体分离。所以，研究还在继续中……

我盼望知道更多物种的基因组序列，从而可能指引我们了解建立起我们今天所见的表观遗传过程的一个进化事件顺序表。例如，酿酒酵母没有 RNAi 机制，但许多其他真菌具有，通过研究不同种类之间系统发育的断层，我们也许会发现是什么事件导致酿酒酵母不再“需要”这个系统。

和生物研究的其他领域相比，表观遗传学似乎更多地建立在试图解释意料之外的现象，如从 H. J. Muller 的位置效应花斑，到 callipyge 表型中的极性超显性（Georges et al. 2004）。试图了解一些非寻常事件的愿望一直吸引着我们，我们为其聪慧灵巧的机制所着迷。这也许可以解释为何这一领域吸引了多于常规的轻松愉快和智慧的头脑。随着我们理解的加深，随着新的、意料之外的表观遗传现象被发现，我觉得这一领域会越来越吸引人。

## 致　谢

感谢我在芝加哥大学和 Fred Hutchinson 癌症研究中心的同事，他们使我对表观遗传学的研究成为一种享受。感谢美国国立健康研究院（National Institutes of Health，NIH）对我研究的资助。

郑　泳　译
朱　冰　审校

## 参 考 文 献

Abraham J., Feldman J., Nasmyth K.A., Strathern J.N., Klar A.J., Broach J.R., and Hicks J.B. 1983. Sites required for position-effect regulation of mating-type information in yeast. *Cold Spring Harbor Symp. Quant. Biol.* **47:** 989–998.

Allfrey V.G., Inoue A., Karn J., Johnson E.M., and Vidali G. 1974. Phosphorylation of DNA-binding nuclear acidic proteins and gene activation in the HeLa cell cycle. *Cold Spring Harbor Symp. Quant. Biol.* **38:** 785–801.

Andrulis E.D., Neiman A.M., Zappulla D.C., and Sternglanz R. 1998. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394:** 592–595.

Andrulis E.D., Zappulla D.C., Ansari A., Perrod S., Laiosa C.V., Gartenberg M.R., and Sternglanz R. 2002. Esc1, a nuclear periphery protein required for Sir4-based plasmid anchoring and partitioning. *Mol. Cell. Biol.* **22:** 8292–8301.

Aparicio O.M. and Gottschling D.E. 1994. Overcoming telomeric silencing: A *trans*-activator competes to establish gene expression in a cell cycle-dependent way. *Genes Dev.* **8:** 1133–1146.

Ariel M., Selig S., Brandeis M., Kitsberg D., Kafri T., Weiss A., Keshet I., Razin A., and Cedar H. 1993. Allele-specific structures in the mouse Igf2-H19 domain. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 307–313.

Bell A., Boyes J., Chung J., Pikaart M., Prioleau M.N., Recillas F., Saitoh N., and Felsenfeld G. 1998. The establishment of active chromatin domains. *Cold Spring Harbor Symp. Quant. Biol.* **63:** 509–514.

Bell S.P., Marahrens Y., Rao H., and Stillman B. 1993. The replicon

model and eukaryotic chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 435–442.

Belote J.M., McKeown M.B., Andrew D.J., Scott T.N., Wolfner M.F., and Baker B.S. 1985. Control of sexual differentiation in *Drosophila* melanogaster. *Cold Spring Harbor Symp. Quant. Biol.* **50:** 605–614.

Beutler E. 1964. Gene inactivation: The distribution of gene products among populations of cells in heterozygous humans. *Cold Spring Harbor Symp. Quant. Biol.* **29:** 261–271.

Bird A.P. 1993. Functions for DNA methylation in vertebrates. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 281–285.

Borst P., Gommers-Ampt J.H., Ligtenberg M.J., Rudenko G., Kieft R., Taylor M.C., Blundell P.A., and van Leeuwen F. 1993. Control of antigenic variation in African trypanosomes. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 105–114.

Brink R.A. 1958. Paramutation at the R locus in maize. *Cold Spring Harbor Symp. Quant. Biol.* **23:** 379–391.

Campbell A. 1981. Some general questions about movable elements and their implications. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 1–9.

Cattanach B.M. and Kirk M. 1985. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315:** 496–498.

Cedar H., Stein R., Gruenbaum Y., Naveh-Many T., Sciaky-Gallili N., and Razin A. 1983. Effect of DNA methylation on gene expression. *Cold Spring Harbor Symp. Quant. Biol.* **47:** 605–609.

Chambon P. 1978. Summary: The molecular biology of the eukaryotic genome is coming of age. *Cold Spring Harbor Symp. Quant. Biol.* **42:** 1209–1234.

Cheng T.H. and Gartenberg M.R. 2000. Yeast heterochromatin is a dynamic structure that requires silencers continuously. *Genes Dev.* **14:** 452–463.

Cheutin T., McNairn A.J., Jenuwein T., Gilbert D.M., Singh P.B., and Misteli T. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299:** 721–725.

Clark D., Reitman M., Studitsky V., Chung J., Westphal H., Lee E., and Felsenfeld G. 1993. Chromatin structure of transcriptionally active genes. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 1–6.

Cockell M., Gotta M., Palladino F., Martin S.G., and Gasser S.M. 1998. Targeting Sir proteins to sites of action: A general mechanism for regulated repression. *Cold Spring Harbor Symp. Quant. Biol.* **63:** 401–412.

Cuthbert G.L., Daujat S., Snowden A.W., Erdjument-Bromage H., Hagiwara T., Yamada M., Schneider R., Gregory P.D., Tempst P., Bannister A.J., and Kouzarides T. 2004. Histone deimination antagonizes arginine methylation. *Cell* **118:** 545–553.

Dion M.F., Altschuler S.J., Wu L.F., and Rando O.J. 2005. Genomic characterization reveals a simple histone H4 acetylation code. *Proc. Natl. Acad. Sci.* **102:** 5308–5309.

Doerfler W., Kruczek I., Eick D., Vardimon L., and Kron B. 1983. DNA methylation and gene activity: The adenovirus system as a model. *Cold Spring Harbor Symp. Quant. Biol.* **47:** 593–603.

Feinberg A.P., Kalikin L.M., Johnson L.A., and Thompson J.S. 1994. Loss of imprinting in human cancer. *Cold Spring Harbor Symp. Quant. Biol.* **59:** 357–364.

Fox C.A., Loo S., Rivier D.H., Foss M.A., and Rine J. 1993. A transcriptional silencer as a specialized origin of replication that establishes functional domains of chromatin. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 443–455.

Frankel J. 1990. Positional order and cellular handedness. *J. Cell Sci.* **97:** 205–211.

Gartler S.M., and Linder D. 1964. Selection In mammalian mosaic cell populations. *Cold Spring Harbor Symp. Quant. Biol.* **29:** 253–260.

Gasser S.M., Hediger F., Taddei A., Neumann F.R., and Gartenberg M.R. 2004. The function of telomere clustering in yeast: The circe effect. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 327–337.

Georges M., Charlier C., Smit M., Davis E., Shay T., Tordoir X., Takeda H., Caiment F., and Cockett N. 2004. Toward molecular understanding of polar overdominance at the ovine callipyge locus. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 477–483.

Goldschmidt R.B. 1951. The theory of the gene: Chromosomes and genes. *Cold Spring Harbor Symp. Quant. Biol.* **16:** 1–11.

Haber J.E., Weiffenbach B., Rogers D.T., McCusker J., and Rowe L.B. 1981. Chromosomal rearrangements accompanying yeast mating-type switching: Evidence for a gene-conversion model. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 991–1002.

Haig D. 2004. The (dual) origin of epigenetics. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 67.

Henikoff S. 1990. Position-effect variegation after 60 years. *Trends Genet.* **6:** 422–426.

Henikoff S., McKittrick E., and Ahmad K. 2004. Epigenetics, histone H3 variants, and the inheritance of chromatin states. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 235–243.

Hernday A.D., Braaten B.A., and Low D.A. 2003. The mechanism by which DNA adenine methylase and PapI activate the pap epigenetic switch. *Mol. Cell* **12:** 947–957.

Hodgkin J., Doniach T., and Shen M. 1985. The sex determination pathway in the nematode *Caenorhabditis elegans:* Variations on a theme. *Cold Spring Harbor Symp. Quant. Biol.* **50:** 585–593.

Imai S., Johnson F.B., Marciniak R.A., McVey M., Park P.U., and Guarente L. 2000. Sir2: An NAD-dependent histone deacetylase that connects chromatin silencing, metabolism, and aging. *Cold Spring Harbor Symp. Quant. Biol.* **65:** 297–302.

Jenuwein T. and Allis C.D. 2001. Translating the histone code. *Science* **293:** 1074–1080.

Klar A.J., and Bonaduce M.J. 1993. The mechanism of fission yeast mating-type interconversion: Evidence for two types of epigenetically inherited chromosomal imprinted events. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 457–465.

Klar A.J., Hicks J.B., and Strathern J.N. 1981. Irregular transpositions of mating-type genes in yeast. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 983–990.

Kubicek S. and Jenuwein T. 2004. A crack in histone lysine methylation. *Cell* **119:** 903–906.

La Volpe A., Taggart M., Macleod D., and Bird A. 1983. Coupled demethylation of sites in a conserved sequence of *Xenopus* ribosomal DNA. *Cold Spring Harbor Symp. Quant. Biol.* **47:** 585–592.

Laroche T., Martin S.G., Gotta M., Gorham H.C., Pryde F.E., Louis E.J., and Gasser S.M. 1998. Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. *Curr. Biol.* **8:** 653–656.

Laurenson P. and Rine J. 1992. Silencers, silencing, and heritable transcriptional states. *Microbiol. Rev.* **56:** 543–560.

Lewis E.B. 1985. Regulation of the genes of the bithorax complex in *Drosophila. Cold Spring Harbor Symp. Quant. Biol.* **50:** 155–164.

Li E., Beard C., Forster A.C., Bestor T.H., and Jaenisch R. 1993. DNA methylation, genomic imprinting, and mammalian development. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 297–305.

Locke J., Kotarski M.A., and Tartof K.D. 1988. Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **120:** 181–198.

Lolle S.J., Victor J.L., Young J.M., and Pruitt R.E. 2005. Genome-wide non-mendelian inheritance of extra-genomic information in *Arabidopsis. Nature* **434:** 505–509.

Losick R. 1998. Summary: Three decades after sigma. *Cold Spring Harbor Symp. Quant. Biol.* **63:** 653–666.

Louie A.J., Candido E.P., and Dixon G.H. 1974. Enzymatic modifications and their possible roles in regulating the binding of basic proteins to DNA and in controlling chromosomal structure. *Cold Spring Harbor Symp. Quant. Biol.* **38:** 803–819.

Lyon M.F. 1961. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190:** 372–373.

———. 1993. Epigenetic inheritance in mammals. *Trends Genet.* **9:** 123–128.

Maine E.M., Salz H.K., Schedl P., and Cline T.W. 1985. Sex-lethal, a link between sex determination and sexual differentiation in *Drosophila melanogaster. Cold Spring Harbor Symp. Quant. Biol.* **50:** 595–604.

Martin-Morris L.E., Loughney K., Kershisnik E.O., Poortinga G., and Henikoff S. 1993. Characterization of sequences responsible for

*trans*-inactivation of the *Drosophila brown* gene. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 577–584.

McClintock B. 1951. Chromosome organization and genic expression. *Cold Spring Harbor Symp. Quant. Biol.* **16:** 13–47.

———. 1956. Controlling elements and the gene. *Cold Spring Harbor Symp. Quant. Biol.* **21:** 197–216.

Megee P.C., Morgan B.A., and Smith M.M. 1995. Histone H4 and the maintenance of genome integrity. *Genes Dev.* **9:** 1716–1727.

Millar C.B., Kurdistani S.K., and Grunstein M. 2004. Acetylation of yeast histone H4 lysine 16: A switch for protein interactions in heterochromatin and euchromatin. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 193–200.

Mirkovitch J., Gasser S.M., and Laemmli U.K. 1987. Relation of chromosome structure and gene expression. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **317:** 563–574.

Mostoslavsky R., Kirillov A., Ji Y.H., Goldmit M., Holzmann M., Wirth T., Cedar H., and Bergman Y. 1999. Demethylation and the establishment of κ allelic exclusion. *Cold Spring Harbor Symp. Quant. Biol.* **64:** 197–206.

Muller H.J. 1941. Induced mutations in *Drosophila*. *Cold Spring Harbor Symp. Quant. Biol.* **9:** 151–167.

Nance W.E. 1964. Genetic tests with a sex-linked marker: Glucose-6-phosphate dehydrogenase. *Cold Spring Harbor Symp. Quant. Biol.* **29:** 415–425.

Nanney D.L. 1958. Epigenetic factors affecting mating type expression in certain ciliates. *Cold Spring Harbor Symp. Quant. Biol.* **23:** 327–335.

Nasmyth K.A., Tatchell K., Hall B.D., Astell C., and Smith M. 1981. Physical analysis of mating-type loci in *Saccharomyces cerevisiae*. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 961–981.

Palladino F., Laroche T., Gilson E., Pillus L., and Gasser S.M. 1993. The positioning of yeast telomeres depends on SIR3, SIR4, and the integrity of the nuclear membrane. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 733–746.

Pannetier C., Hu–Li J., and Paul W.E. 1999. Bias in the expression of IL-4 alleles: The use of T cells from a GFP knock-in mouse. *Cold Spring Harbor Symp. Quant. Biol.* **64:** 599–602.

Peacock W.J., Brutlag D., Goldring E., Appels R., Hinton C.W., and Lindsley D.L. 1974. The organization of highly repeated DNA sequences in *Drosophila melanogaster* chromosomes. *Cold Spring Harbor Symp. Quant. Biol.* **38:** 405–416.

Pillus L. and Rine J. 1989. Epigenetic inheritance of transcriptional states in *S. cerevisiae*. *Cell* **59:** 637–647.

Ptashne M. 2004. *A genetic switch: Phage lambda revisited*, 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Renauld H., Aparicio O.M., Zierath P.D., Billington B.L., Chhablani S.K., and Gottschling D.E. 1993. Silent domains are assembled continuously from the telomere and are defined by promoter distance and strength, and by SIR3 dosage. *Genes Dev.* **7:** 1133–1145.

Rine J., Jensen R., Hagen D., Blair L., and Herskowitz I. 1981. Pattern of switching and fate of the replaced cassette in yeast mating-type interconversion. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 951–960.

Rubin G.M. 1985. Summary. *Cold Spring Harbor Symp. Quant. Biol.* **50:** 905–908.

Rubin G.M., Hazelrigg T., Karess R.E., Laski F.A., Laverty T., Levis R., Rio D.C., Spencer F.A., and Zuker C.S. 1985. Germ line specificity of P-element transposition and some novel patterns of expression of transduced copies of the white gene. *Cold Spring Harbor Symp. Quant. Biol.* **50:** 329–335.

Rudkin G.T. and Tartof K.D. 1974. Repetitive DNA in polytene chromosomes of *Drosophila melanogaster*. *Cold Spring Harbor Symp. Quant. Biol.* **38:** 397–403.

Schubeler D., MacAlpine D.M., Scalzo D., Wirbelauer C., Kooperberg C., van Leeuwen F., Gottschling D.E., O'Neill L.P., Turner B.M., Delrow J., et al. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **18:** 1263–1271.

Schultz J. 1956. The relation of the heterochromatic chromosome regions to the nucleic acids of the cell. *Cold Spring Harbor Symp. Quant. Biol.* **21:** 307–328.

Selker E.U., Richardson G.A., Garrett-Engele P.W., Singer M.J., and Miao V. 1993. Dissection of the signal for DNA methylation in the ζ–η region of *Neurospora*. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 323–329.

Shapiro L.J. and Mohandas T. 1983. DNA methylation and the control of gene expression on the human X chromosome. *Cold Spring Harbor Symp. Quant. Biol.* **47:** 631–637.

Shi Y., Lan F., Matson C., Mulligan P., Whetstine J.R., Cole P.A., and Casero R.A. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119:** 941–953.

Si K., Lindquist S., and Kandel E. 2004. A possible epigenetic mechanism for the persistence of memory. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 497–498.

Solter D., Aronson J., Gilbert S.F., and McGrath J. 1985. Nuclear transfer in mouse embryos: Activation of the embryonic genome. *Cold Spring Harbor Symp. Quant. Biol.* **50:** 45–50.

Swift H. 1974. The organization of genetic material in eukaryotes: Progress and prospects. *Cold Spring Harbor Symp. Quant. Biol.* **38:** 963–979.

Thompson J.S., Hecht A., and Grunstein M. 1993. Histones and the regulation of heterochromatin in yeast. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 247–256.

Tilghman S.M., Bartolomei M.S., Webber A.L., Brunkow M.E., Saam J., Leighton P.A., Pfeifer K., and Zemel S. 1993. Parental imprinting of the H19 and Igf2 genes in the mouse. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 287–295.

Vazquez J., Farkas G., Gaszner M., Udvardy A., Muller M., Hagstrom K., Gyurkovics H., Sipos L., Gausz J., Galloni M., et al. 1993. Genetic and molecular analysis of chromatin domains. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 45–54.

Wade P.A., Jones P.L., Vermaak D., Veenstra G.J., Imhof A., Sera T., Tse C., Ge H., Shi Y.B., Hansen J.C., and Wolffe A.P. 1998. Histone deacetylase directs the dominant silencing of transcription in chromatin: Association with MeCP2 and the Mi-2 chromodomain SWI/SNF ATPase. *Cold Spring Harbor Symp. Quant. Biol.* **63:** 435–445.

Wang Y., Wysocka J., Perlin J.R., Leonelli L., Allis C.D., and Coonrod S.A. 2004. Linking covalent histone modifications to epigenetics: The rigidity and plasticity of the marks. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 161–169.

Weintraub H. 1974. The assembly of newly replicated DNA into chromatin. *Cold Spring Harbor Symp. Quant. Biol.* **38:** 247–256.

———. 1993. Summary: Genetic tinkering local problems, local solutions. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 819–836.

Weintraub H., Flint S.J., Leffak I.M., Groudine M., and Grainger R.M. 1978. The generation and propagation of variegated chromosome structures. *Cold Spring Harbor Symp. Quant. Biol.* **42:** 401–407.

Wickner R.B., Edskes H.K., Ross E.D., Pierce M.M., Baxa U., Brachmann A., and Shewmaker F. 2004a. Prion genetics: New rules for a new kind of gene. *Annu. Rev. Genet.* **38:** 681–707.

Wickner R.B., Edskes H.K., Ross E.D., Pierce M.M., Shewmaker F., Baxa U., and Brachmann A. 2004b. Prions of yeast are genes made of protein: Amyloids and enzymes. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 489–496.

Willard H.F., Brown C.J., Carrel L., Hendrich B., and Miller A.P. 1993. Epigenetic and chromosomal control of gene expression: Molecular and genetic analysis of X chromosome inactivation. *Cold Spring Harbor Symp. Quant. Biol.* **58:** 315–322.

Wood W.B., Meneely P., Schedin P., and Donahue L. 1985. Aspects of dosage compensation and sex determination in *Caenorhabditis elegans*. *Cold Spring Harbor Symp. Quant. Biol.* **50:** 575–583.

Yarmolinsky M.B. 1981. Summary. *Cold Spring Harbor Symp. Quant. Biol.* **45:** 1009–1015.

Zheng C., and Hayes J.J. 2003. Structures and interactions of the core histone tail domains. *Biopolymers* **68:** 539–546.

# 第 2 章

# 表观遗传学发展简史

Gary Felsenfeld

*National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0540*

## 本章目录

1. 引言
2. 源于遗传与发育的线索
3. 个体体细胞 DNA 是相同的
4. DNA 甲基化的作用
5. 染色质的作用
6. 所有机制是相互关联的
参考文献

## 1. 引　言

表观遗传学的发展与进化研究和发育研究密切相关。在过去的 50 年间，随着人们对真核基因表达调控分子机制研究的深入，“表观遗传学”自身的含义也在不断更新，目前通行的定义简要概括了其研究内容，即“研究在有丝分裂及减数分裂过程中无法用 DNA 序列改变来解释的基因功能的可继承性改变”（Riggs et al. 1996）。20 世纪 50 年代，表观遗传学的定义还相当的宽泛，它涵盖了从受精卵起始到个体成熟过程中的所有发育事件，即从遗传物质开始到个体成型所有的调控过程（Waddington 1953），这种概念发源于 19 世纪后期的细胞生物学及胚胎生物学研究，这些工作为研究基因与发育之间的关系奠定了基础。从那时起的相当长时间里，胚胎生物学家对究竟是什么物质承担个体发育的“编程”和对这种物质的特性及定位等方面的理解存在争议。人们做了大量的细胞及胚胎实验，尽管这些实验设计巧妙但最终解释含糊，依据各自理解的不同，胚胎生物学家们分为两派：一派认为每个细胞含有既定的程序性的要素，它们在发育过程中放大；另一派认为是可溶性物质的化学反应执行了复杂的发育编程，二者的分歧其实是对细胞核和细胞质在发育过程中的重要性的认识不同。1879 年，Flemming 发现了染色体，随后很多研究者，包括 Wilson 和 Boveri 通过实验证实染色体是发育程序的载体，Thomas Hunt Morgan 的果蝇 X 染色体遗传连锁实验为这一观点提供了最具说服力的证据。

从那时起，线性染色体图谱建立迅速推进，果蝇基因被定位在染色体的特定位点（Sturtevant 1913）。然而，问题依旧：染色体中的什么分子承载着遗传信息，它们如何控制发育的进程，它们在细胞分裂中又如何传递？尽管当时已经知道染色体包括核酸和蛋白质，但它们的相对贡献依然不清楚；没人相信核酸能够独自承载发育信息，以前人们争论的问题甚至仍在继续着细胞质对发育事件的可能贡献。果蝇遗传实验证据显示在没有基因改变时表型遗传变化依然存在。当确定 DNA 是遗传信息的基本载体时，争论终于告一段落。所以，重新定义表观遗传学从而区分来自 DNA 序列改变和非 DNA 序列改变引起的可继承的变化显得尤其必要。

## 2. 源于遗传与发育的线索

无论定义怎么变化，从 20 世纪初至今，现代表观遗传学无论是理论观点还是科学数据都在逐步积累。1930 年，H. J. Muller 描述了一种称为“常变位移”（eversporting displacement，“常变”是指高频率的表型变化）的果蝇高频率表型变化突变（Muller 1930）。这些突变涉及染色体转移（translocation）（即位移，displacement），但是即使所有染色质以正确剂量存在，只是排列异常，表型结果也并非都正常。某些个体中，Muller 观察到果蝇有花斑眼。他曾考虑这可能是“不同的眼形成细胞遗传多样性”的结果，但是进一步的遗传学分析使他想到了染色体重排的异常特性，因此他得出结论“染色体上影响各种特征的区域是有某种程度相关的，而不是以个体基因或位置“基因

元件”存在。在此后的 10～20 年间，诸多实验室提供了强有力的证据证实了 *white* 基因移位到异染色质区会产生这样的变异。

在此期间，染色体重排是关注的重点。基因并非完全独立，它们的功能可以受所处基因组中的位置影响，这被多种果蝇突变体充分验证，其中有的突变体果蝇具有不同的花斑眼，有些则涉及基因向常染色质区的转移导致更具普遍性（非花斑型）的位置效应。McClintock 的研究工作（1965）更加清晰地揭示了转座因子在植物遗传中的作用机理。

另一条线索来自对发育过程的研究。在发育过程中，分化的细胞和组织之间存在表型差异，而且差异一旦形成，细胞的特征可以以一种克隆性的方式遗传给子代细胞，尽管可以认为其根源是细胞特异性编程的存在，但对它究竟是以何种方式向子代细胞传递依然鲜有了解。

因此，研究人员提出了大量的设想试图对具体的作用机理加以勾勒：持生物化学观点的研究人员认为，是多种相互依赖的生物化学反应决定并维持了一类细胞的同一性。Delbruck 于 1949 年提出假说认为，如果将生化途径进行简单的两两配对，每一个反应的产物将会是其他途径的中间产物或抑制剂，依此建立的系统就可以在一个或两个稳定状态之间转换。随后的发现，如细菌 *lac* 操纵子（Novick and Weiner 1957）机理、噬菌体溶原性与裂解性状态的转换（Ptashne 1992）是该理论的直接例证。可以预见，在真核生物中存在功能等价的模型。在胚胎发育过程中，细胞核与细胞质分别对分化状态的转变贡献究竟有多大，自然成了最感兴趣的话题，同时也是争议的焦点所在，然而事实却是，自稳定生化途径在细胞分裂中必须得以维持。草履虫及其他纤毛虫则表现出另外一类表观遗传转变，纤毛形态在个体之间存在差异而且可以克隆性继承（Beisson and Sonneborn 1965）。通过显微外科技术改变其皮层形态后，子代也获得了这种被改变的皮层形态。有人认为线虫中也存在类似的现象，细胞组分受到特定的细胞质决定因子的影响通过一定的方式在细胞分裂期间进行传递（Grimes and Aufderheide 1991）。

### 3. 个体体细胞 DNA 是相同的

尽管染色体存在于所有的体细胞，但当时并不确定所有的体细胞保留了受精卵中所有的 DNA。直到 Avery、MacLeod 和 McCarty 的工作（1944）以及 Hershey 和 Chase（1952）的工作发表后，人们才更为清楚地认识到，DNA 分子可能承载着遗传信息，他们的观点被 Watson 和 Crick 在 1953 年通过解析 DNA 的结构得以确证。Briggs 和 King 在豹蛙（*Rana pipiens*）中的研究（1952）以及 Laskey 和 Gurdon 在爪蟾（*Xenopus*）中的研究（1970）得到同样的结果，那就是：将来自早期胚胎细胞的核植入去核的卵母细胞可以使其发育成胚胎。因此 Laskey 和 Gurdon 认为：成年动物的体细胞同时拥有其自身生长和发育所必需的基因以外的基因。在这篇文章中，他们进一步表明一个体细胞核 DNA 一旦被植入去核的卵细胞，就具有诱导胚胎发育的潜能。现在人们已经清楚，发育的进程以及体细胞所有组成成分表达的特异性其实受控于某些信号，而这些信号并非来自生殖细胞 DNA 序列的缺失或突变。

当然，体细胞 DNA 可以表现得与生殖细胞有所不同，其结果是产生了细胞表型，Barbara McClintock 以及其他植物遗传学家的工作表明转座子可以改变体细胞的表达模式；抗体多样性的产生是因为体细胞 DNA 重排，它与早期 Muller 所描述的位置效应变异一样，可以看做是一个表观遗传事件。然而，近年来表观遗传学的很多工作都集中在研究没有 DNA 重排发生时的系统，这包括 DNA 碱基的修饰以及与 DNA 形成复合体的核蛋白的修饰。

## 4. DNA 甲基化的作用

X 染色体失活为这类表观遗传机理的研究提供了早期模型（Ohno et al. 1959；Lyon 1961），体细胞中哪条 X 染色体的失活是随机选择的，DNA 的序列本身没有发生改变。为了（部分地）解释这类失活作用的机理，Riggs（1975）、Holliday 和 Pugh（1975）等提出将 DNA 甲基化作为表观遗传学标志。该模型的核心是：甲基化位点位于回文序列区；特定的甲基化酶区分并分别催化未修饰双链 DNA 和已修饰单链 DNA 的甲基化。该观点认为，首次甲基化比再次甲基化要困难得多，一旦第一链 DNA 被修饰，互补链的相同回文位点也很快甲基化，亲本链上的甲基化标志可以随着复制的进行拷贝到子代链，因此甲基化状态正确地传递给了子代。动物的甲基化位点主要是发生在 CpG 序列（Doskocil and Sorm 1962），因此，Bird 利用甲基化敏感的限制性内切酶来检测甲基化状态，结果发现内源 CpG 位点的甲基化状态有两种：即要么完全未甲基化，要么完全甲基化（Bird 1978；Bird and Southern 1978）。早期模型的预测从而得以证实并阐明了甲基化标志通过半保留复制得以传递的机理。

随后的几年中，大量工作集中在以下几个方面：内源 DNA 甲基化的模式及其通过生殖细胞的可能传递方式，DNA 甲基化在基因表达沉默中的作用，完全非甲基化位点的首次甲基化起始或抑制的机制以及鉴定那些催化再次甲基化和维持甲基化的酶。虽然在脊椎动物中发现的大部分 DNA 甲基化与重复序列和逆转录病毒相关，并可以使这些序列维持在永久沉默的状态，但在许多情况下，这种修饰毫无疑问为基因活性的表观遗传学传递提供了基础。印记位点（Cattanach and Kirk 1985），如鼠或人类 *Igf2/H19* 位点，其中一个等位基因 DNA 被甲基化修饰，从而控制了二个等位基因的表达（Bell and Felsenfeld 2000；Hark et al. 2000），这显然不是表观遗传学信息传递的唯一机制，例如，位置效应变异现象很早以前就在果蝇当中被发现，然而果蝇的 DNA 甲基化水平却非常低。随后，遗传学家确定了果蝇的聚梳（*Polycomb*）和三胸（*Trithorax*）类基因，发现它们引起发育过程中一系列基因活性状态的永久“锁定”：或者持久关闭或者持久开启。这种状态在细胞分裂过程中稳定传递，暗示一种表观遗传机制存在于其中。

## 5. 染色质的作用

人们早就认识到，在真核生物的细胞核中结合于 DNA 上的蛋白质，尤其是组蛋白，可能参与了 DNA 的修饰。早在 DNA 甲基化研究开展之前，Stedman 和 Stedman

(1950) 就提出，组蛋白是基因表达的广谱抑制因子。他们认为，因为生物体的所有体细胞有相同数目的染色体，它们有相同的遗传学因子（虽然多年后才被证实）。即使认识到组蛋白修饰的奥妙是距其很久以后的事情，Stedman 二人依然提出了自己的假设，他认为在一个生物体中，各种不同的细胞具有不同的组蛋白，从而使表型产生明显的差异。组蛋白的确可以减少转录水平使之远远低于原核生物失活基因的转录水平。随后的工作研究了染色质作为转录模板的功能，并检验了这种功能是否以细胞特异性的方式受到限制。在 1963 年的一篇论文中，Bonner (Bonner et al. 1963) 从豌豆类植物的产球蛋白组织中提取染色质，并发现在体外体系中，当加入大肠杆菌 RNA 聚合酶，产生的转录物得到翻译，球蛋白可被检测到，并且该结果仅特异性地发生在该组织中。随着杂交方法的出现，这种体外实验转录物可被检测 (Paul and Gilmour 1968) 并且特异地来自于获得该染色质的特定的组织。其他结果表明，这种特异性显示到达转录起始位置受到抑制 (Cedar and Felsenfeld 1973)。然而，曾有一段时期人们普遍认为组蛋白是抑制蛋白，可被动地沉默基因的表达，以这种观点看来，激活基因就只意味着去除组蛋白；一旦组蛋白被去除，转录就会像在原核生物中那样进行。不过，有一些证据表明，真核细胞中不存在开放 DNA 的伸展区域 (Clark and Felsenfeld 1971)，即使 DNA 裸露的模型是正确的，但是究竟哪些区域覆盖的组蛋白应该首先被去除，这个结论如何下依然是个问题。

早在 1964 年，Allfrey 等 (1964) 推测组蛋白乙酰化可能与基因激活相关，“活跃的”染色质无须组蛋白脱离 DNA。随后的十年，人们对研究组蛋白修饰与基因表达之间的关系表现出极大的兴趣，乙酰化之外的修饰（甲基化和磷酸化）也随之被确认，但它们的功能意义仍然不清楚。Kornberg 和 Thomas (1974) 解析了基本的染色质单元——核小体的结构之后，解决这个问题就容易了很多。核小体的晶体结构测定的分辨率，首先是 7Å，然后是 2.8Å，结果不仅提供了重要的结构信息，更显示出组蛋白 N 端尾巴部分伸出 DNA -蛋白质的八聚体核心，使之更容易被修饰 (Richmond et al. 1984; Luger et al. 1997)。从 20 世纪 80 年代开始的若干年内，Grunstein 及合作者们 (Wallis et al. 1980; Durrin et al. 1991) 应用酵母进行遗传分析，结果表明组蛋白 N 端尾巴，对于基因表达调控以及沉默染色质域的形成是必需的。

而对其详细的机制研究始于 Allis (Brownell et al. 1996) 的研究证据，即四膜虫的组蛋白乙酰转移酶与酵母转录调控蛋白 Gcn5 同源，这提供了直接的证据把组蛋白乙酰化与基因表达调控联系到一起。自那时以来，组蛋白修饰的研究变得炙手可热，同时人们对已知修饰的作用机理也开始进行重新评估。

但是，这些修饰位点在体内如何被选择仍然有待回答。例如，Gal4-VP16 已被证明能够以 ATP 依赖的方式从重组的染色质模板激活转录，激活伴随核小体重新定位，这被认为是使启动子易于接近的关键事件 (Pazin et al. 1994)。更深入细致地了解这些发现的意义就需要鉴定 ATP 依赖性核小体重塑复合物如 SWI/SNF 和 NURF (Peterson and herskowitz 1992; Tsukiyama and Wu 1995) 以及认识到组蛋白修饰和核小体重塑复合体二者都是染色质模板转录所必需的。

基因活性状态的信息如何通过这种机制在细胞分裂中传递仍然不清楚；因此其在表

观遗传学信息传递中的作用也不清楚。随后的重大进步是认识到组蛋白以其修饰所特有的专一性方式在其周围募集那些可以影响染色质的局部结构和功能状态的蛋白质。例如，组蛋白 H3K9 甲基化导致异染色质蛋白 HP1 在其周围富集（Bannister et al. 2001；Lachner et al. 2001；Nakayama et al. 2001），HP1 进而可以募集甲基化酶（Suv39 h1），从而使这个区域逐步甲基化从而沉默（图 2-1a）。这同时也为该状态如何通过复制周期被传递并且继续存在提供了一个合理的解释（图 2-1b）。基因活跃状态也存在类似的延伸机制，其中涉及组蛋白 H3K4 甲基化和 Trithorax 蛋白的参与（Wysocka et al. 2005）。

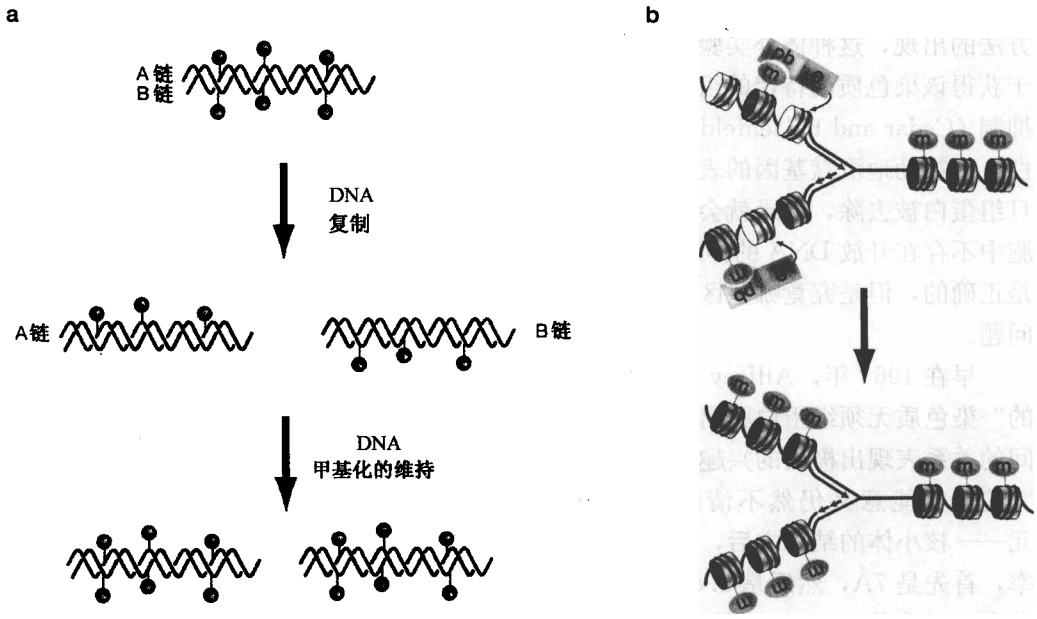



图 2-1 DNA 复制期间维持 DNA 甲基化和组蛋白修饰模式的机制

（a）DNA 复制期间维持 DNA 甲基化模式的机制。复制期间，一条 DNA 链的 CpG 或者 CpXpG 残基上存在特异的甲基化模式，开始配对合成另一条新的未甲基化的 DNA 链。每条 DNA 链上的 CpG 会有相应的 CpG 在另一条链上存在。此时 DNA 甲基转移酶会识别半甲基化的位点并甲基化新链上的胞嘧啶，以维持甲基化的模式不被破坏。（b）复制期间维持组蛋白修饰的基本机制。修饰的组蛋白尾部（m）与含有特异的修饰结合位点的蛋白结合子（pb）结合；反之 pb 也含有与特异的进行组蛋白修饰的酶（e）的结合位点。酶（e）进而能修饰相邻的核小体。复制期间新插入的组蛋白与亲本组蛋白是散在分布的因此获得了亲本的修饰。在细胞周期的任何阶段相似的机制使得组蛋白修饰从一个修饰的区域向未修饰的区域延伸。

延伸机制也可以依赖于组蛋白变体，而不仅仅是被修饰的组蛋白（Ahmad and Henikoff 2002；McKittrick et al. 2004）。组蛋白 H3 只有通过 DNA 复制才能结合到染色质中。其变体 H3.3 与 H3 有四个氨基酸不同，是以非依赖复制的方式结合到核小体中，而且它往往会聚积到富含“活跃的”组蛋白修饰的激活染色质区域（McKittrick et al. 2004）。H3.3 的存在足以保持活化状态，虽然在复制之后它被稀释到一半，但仍足以继续保持活性状态，随后的转录导致 H3.3 核小体替换原有的 H3 核小体而得以补充，因此该活化状态在子代中得到永久保持。

## 6. 所有机制是相互关联的

这些模型最终导致了组蛋白修饰、组蛋白变体以及特定基因的活化与表观遗传学之间的联系的组建，这无疑还有大量的工作去做。这些机制提供给我们一些关于异染色质状态如何保持的观点，可是它们却不能解释沉默染色质结构首次是如何建立的。直到近来，人们才清楚 RNA 转录产物，尤其是来自重复序列的转录产物，在诸如 Dicer、Argonaute 和 RNA 依赖的 RNA 聚合酶的加工下得到小分子 RNA，这些 RNA 随后被招募到与其互补的同源 DNA 位点作为包含有 Polycomb 类蛋白的复合体的一部分，启动异染色质的形成。有证据显示，一些异染色质区域的维持需要同样的机制。从某种意义上来看，这些稳定循环反应机制其实可以看做是 50 年前 Delbrück 模型的翻版，即稳定的生物化学循环维持着有机体的状态。

我们现在知道，不计其数的表观遗传学机制在有机体中作用的例子。除许多位点的基因印记以及上面提到的等位基因特异性失活和 X 染色体随机失活之外，位于同一染色体上的免疫球蛋白基因的重排被选择性地抑制以及在嗅觉神经元中单一气味受体基因表达也属于表观遗传现象（Chess et al. 1994；Shykind et al. 2004）。在果蝇中，Polycomb 类基因主要负责建立沉默染色质结构域，并在随后所有的细胞分裂中加以维持。植物的副突变也是因为表观遗传的改变，一个等位基因可以引起同源的等位基因的表达发生可继承的变化（Stam et al. 2002），这是一个既可以减数分裂又可以有丝分裂继承的表观遗传状态的例子，在植物中广泛存在而在动物中极少存在（Jorgensen 1993）。上述机制的很多证据来自于对酿酒酵母（*Schizosaccharomyces pombe*）中接合型基因座沉默和着丝粒序列沉默的研究工作（Hall et al. 2002）。复杂的着丝粒（如果蝇和人），其凝聚染色质结构特性已被证明可通过着丝粒相关蛋白而不是 DNA 序列进行转移。在所有这些情况下，DNA 序列保持完整，但是它的表达能力被抑制，这很可能是由 DNA 甲基化、组蛋白修饰，或两者共同介导的，而且在一些情况下被证实的确如此。最后，上述关于草履虫的表观遗传“模式”可以扩展到朊粒，这种蛋白可以使其交替折叠的状态得到维持并传递给子代细胞。

尽管我们这里所讲是一个个顺序发生的故事，但确切地说它们其实是一连串试图去定义并解释表观遗传学现象的事件，它们平行发生而且有所重叠。表观遗传学的定义已经改变，但是早期科学家们有关发育机制的问题依然存在，当代表观遗传学的核心仍然是这些重大问题。自 Muller 描述位置效应花斑，已经过去了 70 年，循着表型的观察，经过细致的遗传学研究，再到近来在分子水平上的解析，这是一个可喜的进步。随着理解的深入，表观遗传学的机制可以应用到相当一部分表型复杂的有机体，而事实却是，我们往往被第一眼看起来就很有趣的现象所迷惑而将主要问题边缘化，从而导致研究重心的偏离，然而我们花了很长时间才认识到这个问题。

常建峰　译
王大亮　校
孙方霖　审校

# 参考文献


Ahmad K. and Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9:** 1191–1200.

Allfrey V.G., Faulkner R., and Mirsky A.E. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci.* **51:** 786–794.

Avery O.T., MacLeod C.M., and McCarty M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* **79:** 137–158.

Bannister A., Zegerman P., Partridge J., Miska E., Thomas J., Allshire R., and Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120–124.

Beisson J. and Sonneborn T.M. 1965. Cytoplasmic inheritance of the organization of the cell cortex in *Paramecium aurelia*. *Proc. Natl. Acad. Sci.* **53:** 275–282.

Bell A.C. and Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405:** 482–485.

Bird A.P. 1978. Use of restriction enzymes to study eukaryotic DNA methylation. II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J. Mol. Biol.* **118:** 49–60.

Bird A.P. and Southern E.M. 1978. Use of restriction enzymes to study eukaryotic DNA methylation. I. The methylation pattern in ribosomal DNA from *Xenopus laevis*. *J. Mol. Biol.* **118:** 27–47.

Bonner J., Huang R.C., and Gilden R.V. 1963. Chromosomally directed protein synthesis. *Proc. Natl. Acad. Sci.* **50:** 893–900.

Briggs R. and King T.J. 1952. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proc. Natl. Acad. Sci.* **38:** 455–463.

Brownell J.E., Zhou J., Ranalli T., Kobayashi R., Edmondson D.G., Roth S.Y., and Allis C.D. 1996. Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84:** 843–851.

Cattanach B.M. and Kirk M. 1985. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315:** 496–498.

Cedar H. and Felsenfeld G. 1973. Transcription of chromatin in vitro. *J. Mol. Biol.* **77:** 237–254.

Chess A., Simon I., Cedar H., and Axel R. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78:** 823–834.

Clark R.J. and Felsenfeld G. 1971. Structure of chromatin. *Nat. New Biol.* **229:** 101–106.

Doskocil J. and Sorm F. 1962. Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. *Biochim. Biophys. Acta* **55:** 953–959.

Durrin L.K., Mann R.K., Kayne P.S., and Grunstein M. 1991. Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* **65:** 1023–1031.

Grimes G.W. and Aufderheide K.J. 1991. Cellular aspects of pattern formation: The problem of assembly. *Monogr. Dev. Biol.* **22:** 1–94.

Hall I.M., Shankaranarayana G.D., Noma K., Ayoub N., Cohen A., and Grewal S.I. 2002. Establishment and maintenance of a heterochromatin domain. *Science* **297:** 2215–2218.

Hannah A. 1951. Localization and function of heterochromatin in *Drosophila melanogaster*. *Adv. Genet.* **4:** 87–125.

Hark A.T., Schoenherr C.J., Katz D.J., Ingram R.S., Levorse J.M., and Tilghman S.M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405:** 486–489.

Hershey A.D. and Chase M. 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* **36:** 39–56.

Holliday R. and Pugh J.E. 1975. DNA modification mechanisms and gene activity during development. *Science* **187:** 226–232.

Jablonka E. and Lamb M.J. 1995. *Epigenetic inheritance and evolution: The Lamarckian dimension*. Oxford University Press, New York, p. 82.

Jorgensen R. 1993. The germinal inheritance of epigenetic information in plants. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **339:** 173–181.

Kornberg R.D. and Thomas J.O. 1974. Chromatin structure; oligomers of the histones. *Science* **184:** 865–868.

Lachner M., O'Carroll D., Rea S., Mechtler K., and Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410:** 116–120.

Laskey R.A. and Gurdon J.B. 1970. Genetic content of adult somatic cells tested by nuclear transplantation from cultured cells. *Nature* **228:** 1332–1334.

Luger K., Mader A.W., Richmond R.K., Sargent D.F., and Richmond T.J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389:** 251–260.

Lyon M.F. 1961. Gene action in the X-chromosome of the mouse. *Nature* **190:** 372–373.

McClintock B. 1965. The control of gene action in maize. *Brookhaven Symp. Biol.* **18:** 162–184.

McKittrick E., Gafken P.R., Ahmad K., and Henikoff S. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci.* **101:** 1525–1530.

Morgan T. 1911. An attempt to analyze the constitution of the chromosomes on the basis of sex-linked inheritance in *Drosophila*. *J. Exp. Zool.* **11:** 365–414.

Muller H.J. 1930. Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* **22:** 299–334.

Nakayama J., Rice J.C., Strahl B.D., Allis C.D., and Grewal S.I. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292:** 110–113.

Novick A. and Weiner M. 1957. Enzyme induction as an all-or-none phenomenon. *Proc. Natl. Acad. Sci.* **43:** 553–566.

Ohno S., Kaplan W.D., and Kinosita R. 1959. Formation of the sex chromatin by a single X-chromosome in liver cells of *Rattus norvegicus*. *Exp. Cell Res.* **18:** 415–418.

Paul J. and Gilmour R.S. 1968. Organ-specific restriction of transcription in mammalian chromatin. *J. Mol. Biol.* **34:** 305–316.

Pazin M.J., Kamakaka R.T., and Kadonaga J.T. 1994. ATP-dependent nucleosome reconfiguration and transcriptional activation from preassembled chromatin templates. *Science* **266:** 2007–2011.

Peterson C.L. and Herskowitz I. 1992. Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. *Cell* **68:** 573–583.

Ptashne M. 1992. *A genetic switch: Phage λ and higher organisms*, 2nd edition. Blackwell Science, Malden, Massachusetts and Cell Press, Cambridge, Massachusetts.

Richmond T.J., Finch J.T., Rushton B., Rhodes D., and Klug A. 1984. Structure of the nucleosome core particle at 7 Å resolution. *Nature* **311:** 532–537.

Riggs A.D. 1975. X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* **14:** 9–25.

Riggs A.D. and Porter T.N. 1996. Overview of epigenetic mechanisms. In *Epigenetic mechanisms of gene regulation* (ed. V.E.A. Russo et al.), pp. 29–45. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Riggs A.D., Martienssen R.A., and Russo V.E.A. 1996. Introduction. In *Epigenetic mechanisms of gene regulation* (ed. V.E.A. Russo et al.), pp. 1–4. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shykind B.M., Rohani S.C., O'Donnell S., Nemes A., Mendelsohn M., Sun Y., Axel R., and Barnea G. 2004. Gene switching and the stability of odorant receptor gene choice. *Cell* **117:** 801–815.

Stam M., Belele C., Dorweiler J., and Chandler V. 2002. Differential chromatin structure with a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. *Genes Dev.* **16:** 1906–1918.

Stedman E. and Stedman E. 1950. Cell specificity of histones. *Nature* **166:** 780–781.

Sturtevant A. 1913. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.* **14:** 43–59.

Tsukiyama T. and Wu C. 1995. Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell* **83:** 1011–1020.

Waddington C.H. 1953. Epigenetics and evolution. *Symp. Soc. Exp. Biol.* **7:** 186–199.

Wallis J.W., Hereford L., and Grunstein M. 1980. Histone H2B genes of yeast encode two different proteins. *Cell* **22:** 799–805.

Wysocka J., Swigut T., Milne T. Dou Y., Zhang X., Burlingame A., Roeder R., Brivanlou A., and Allis C.D. 2005. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121:** 859–872.

# 第 3 章

# 概述与概念

C. David Allis[1], Thomas Jenuwein[2], and Danny Reinberg[3]

[1]*The Rockefeller University, New York, New York;*
[2]*Research Institute of Molecular Pathology, Vienna, Austria;*
[3]*UMDNJ-Robert Wood Johnson Medical School, Piscataway, New Jersey*

## 本章目录

1. 遗传学与表观遗传学
2. 研究表观遗传学的模式系统
3. 表观遗传学的界定
4. 染色质模板
5. 染色质的高级结构
6. 常染色质与异染色质的区别
7. 组蛋白修饰与组蛋白密码
8. 染色质重塑复合物与组蛋白变体
9. DNA 甲基化
10. RNA 干扰与 RNA 指导的基因沉默
11. 从单细胞系统到多细胞系统
12. 聚梳和三胸
13. X 染色体失活和兼性异染色质
14. 细胞命运重编程
15. 癌
16. 表观遗传控制的意义
17. 表观遗传研究所面临的重大问题
参考文献

## 概要

过去几年内，人类和许多模式生物的基因组测序工作在生物医学领域及公众中引起了极大的反响。承载人们广为接受的孟德尔遗传定律的基因图谱已随时可供查阅，这无疑打开了人们对人类生物学和疾病深入理解的大门。同时，这些知识为新的治疗策略的产生燃起了希望。但是许多根本性的疑问仍然存在。举例来说，虽然每个细胞拥有相同的遗传信息，它们却遵循着具有高度时空精确性的不同的发育途径。那么，正常的发育过程是如何进行的呢？一个细胞是如何决定何时进行分裂和分化，何时保持其细胞特征性，又如何依照其正常的发育程序进行响应和表达？上述过程中出现错误能够导致疾病，如癌症。这些错误是由于从父母处继承了错误的遗传信息，还是存在着尚未被正确理解并破译的其他层面的调节信息？

在人类中，遗传信息（DNA）被编入 23 对染色体，约 25 000 个基因。这些染色体就像是一个个藏书丰富的图书馆，它们共同指导着人的整个发育过程。我们的基因组 DNA 序列由大约 $3\times10^9$ 个碱基组成，它们被简写成 A、C、G 和 T 四个字母（碱基），组成了能够被很好识别的单词（基因）、句子、章节和书。然而，是什么来指导什么时候阅读什么书，依何种顺序阅读，这些问题还不太明了。我们迎接这一巨大挑战很可能揭示在正常和异常的发育过程中，各种细胞事件是如何协调发展的。

高等真核生物中，每个细胞染色体 DNA 分子的总长约 2 米，因此需要将其最大限度地浓缩约 10 000 倍才能被装入细胞核——细胞内遗传物质的存储部位。DNA 分子以组蛋白为核心进行缠绕，产生了蛋白质：DNA 的重复单元，也就是我们所说的染色质，这巧妙地解决了 DNA 的包装问题。然而，由于 DNA 被包装入一个有限的空间范围，因此新的问题产生了。就像是当把大量的书籍摆上图书馆的书架上时，想找到并阅读某一本特定的书变得很困难。因此，必须存在一定的检索系统。染色质作为基因组组装平台，提供了这种检索系统。染色质在结构上并不是均一的，从高度浓缩的染色质纤维状态（异染色质）到相对松散的有利于基因表达的状态（常染色质），它有着不同的包装设计。染色质多聚体可以通过以下方式来引入各种结构上的变化：不同组蛋白的介入（组蛋白变体）；染色质结构的改变（染色质重塑）；一些化学分子对组蛋白的修饰（共价修饰）。另外，甲基基团能够被加到 DNA 模板中的胞嘧啶（C）上，这一过程称为 DNA 甲基化，它能够作为一个停靠点来结合能改变染色质状态的蛋白质，或者影响到原有的组蛋白共价修饰。最近有证据显示非编码 RNA 能够“引导”基因组的特定区域进入高浓缩状态。因此，染色质可以被看做是一个动态的多聚体，它能够对基因组进行检索，还可以响应环境信号，最终决定哪些基因表达，哪些基因不表达。

总体而言，这些调节作用将染色质作为一条组织法则提供给了基因组，称为“表观遗传学”，也就是本书的主题。在某些情况下，表观遗传检索模式可以在细胞分裂过程中被继承，它们为细胞提供了记忆，这样就拓展了遗传密码（DNA）的信息承载量。表观遗传学狭义上可以被定义为通过调整染色质状态，而非 DNA 序列的改变，来实现基因转录调节。

在这篇概述中，我们解释了染色质及表观遗传学的基本概念，并且讨论了为什么表观控制能够帮助我们解决一些一直存在的疑问，例如细胞身份、肿瘤发生、干细胞可塑性、再生及衰老等。当读者阅读后续章节时，我们鼓励读者们能够认识到在众多实验模型中发现的大量的生命现象均有其表观遗传学（非 DNA）基础。在“后基因组”时代，理解表观遗传学的运作机制对于人类生物学及人类疾病具有重要且深远的意义。

## 1. 遗传学与表观遗传学

DNA 双螺旋结构的破译是生命科学领域的一个里程碑式的发现。DNA 是储存遗传信息的主要大分子，生物体通过生殖细胞将储存的遗传信息传递给下一代（Avery et al. 1944）。由于这些和其他的一些发现，现代生物学的“中心法则”出现了。这个法则涵盖了生命必需的遗传信息模板的维持和翻译过程，主要包括：①DNA 以半保留复制的方式进行自我传代；②以遗传密码（DNA）为模板进行 5′端到 3′端的单向转录，产生中间体信使 RNA（mRNA）；③将 mRNA 翻译成从氨基端到羧基端的线性多肽，这是和 DNA 5′端到 3′端的顺序相吻合的。简化一下就是：DNA⟹RNA⟶蛋白质。按照中心法则，RNA 可以通过逆转录然后整合到已经存在的 DNA 中（如通过逆转录病毒和逆转座子）完成与 DNA 的反馈过程。然而，中心法则否认从蛋白质到 DNA 的反馈过程。但有些稀有的蛋白质，例如众所周知的朊粒，可以在没有 DNA 或 RNA 模板的情况下被继承，这看起来似乎有悖于中心法则。显然，这些特殊的自我聚集的蛋白质具有一些和 DNA 类似的性质，包括复制和信息储存的机制（Cohen and Prusiner 1998；Shorter and Lindquist 2005）。另外，越来越多的证据表明我们基因组的相当大一部分是被转录成了“非编码”RNA。这些非编码 RNA（即除了 tRNA、rRNA、snoRNA以外的不编码蛋白质的 RNA）的功能正成为研究的热点，且仅在数量有限的研究中开始变清晰起来。

表观遗传学是在研究与经典孟德尔定律不相符的许多遗传现象过程中逐步发展起来的（见第 1、2 章的历史综述）。经典的孟德尔定律认为表型特征遗传（如豌豆颜色、手指数或者血红蛋白不足）是由于 DNA 序列突变导致的等位基因的差别而造成的。总体而言，突变是定义表型特点的基础，而表型特点又有助于界定物种之间的界线。按照达尔文进化理论，自然选择压力造就了这些界线。这些理论将突变放在了经典遗传学的核心位置。相反，举例而言，非孟德尔定律的继承性（如胚胎生长的异质性、皮肤的嵌合颜色、X 染色体的随机失活、植物的副突变）（图 3-1）可以表现为在同样的细胞核环境中两个等位基因中却只有一个得到表达。重要的是，DNA 序列在这些情况中并没有发生改变。这与我们通常提到的另一种非孟德尔遗传模式是不同的，后者是由于母体线粒体遗传所致（Birky 2001）。

表观遗传学研究的一个重要问题是同一个细胞核内等位基因的选择性调控。是什么机制区分两个完全相同的等位基因？这种区分机制又是如何建立并在连续的细胞传代中维持下去的？是什么原因造成来自同一受精卵（同卵双生，遗传背景完全相同）的双胞胎并不完全相同？表观遗传学被认为是导致外表特征不同的原因之一。环境、饮食和其他潜在的外部因素都可能通过表观遗传影响基因组表达（Klar 2004；见第 23、24 章）。确定什么成分在分子水平被影响，这些成分的改变又是如何影响人类生物学和人类疾病的，将会是未来研究的重大挑战。

这个领域的另一个关键问题是，表观遗传信息对正常发育的贡献有多重要？这些正常的信号通路是如何发生紊乱而导致发育异常和肿瘤发生（即癌症）的？上面提到，同

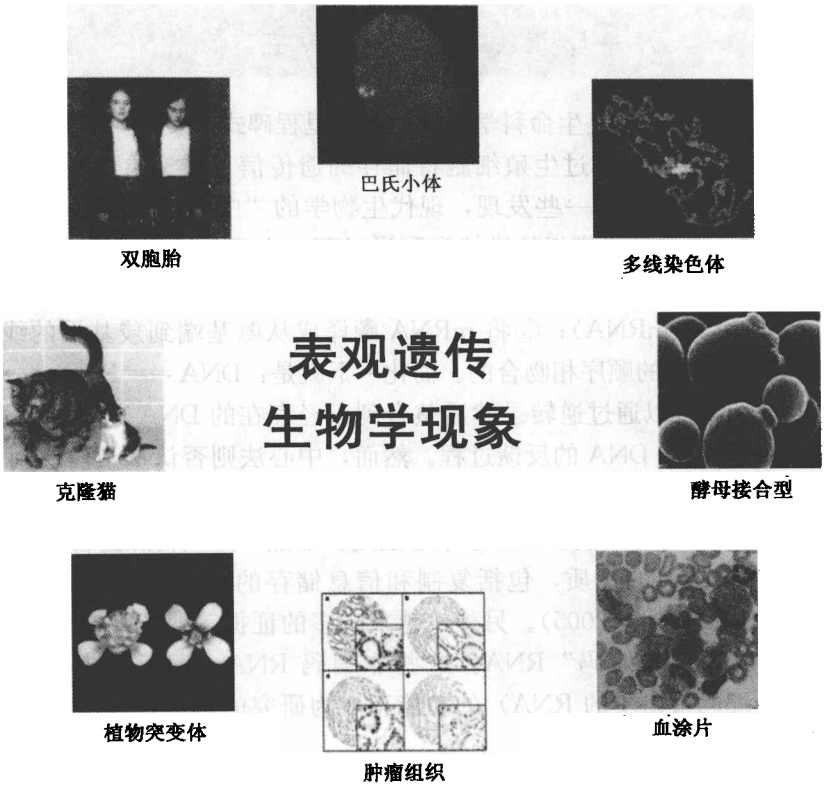



图 3-1 表观遗传表型的生物学示例

大量生物和细胞中的表观遗传类表型的差异都是由非遗传差异造成的。双胞胎：微小的差异主要是表观遗传导致的；(©Randy Harris, New York)；巴氏小体：因表观遗传因素导致哺乳动物雌性细胞中一条 X 染色体失活，该失活染色体在细胞中以致密的异染色质形式存在；多线染色体：果蝇唾腺中的巨大染色体，是研究表观遗传标记与基因表达相关性的理想材料；(引自 Schotta et al. 2003 [©Springer])；酵母接合型：酵母性行为决定于活性的接合位点，此时细胞中其他接合基因的拷贝将沉默，这种基因沉默是表观遗传调控的；(©Alan Wheals, University of Bath)；血涂片：异源细胞虽具有相同表型，但是其不同的功能是由表观遗传决定的；(经允许引自 Christian Sillaber)；肿瘤组织：组织中的亚稳态的细胞（图左）表观遗传标记的表达量上升；(经允许引自 Seligson et al. 2005 [©Macmillan])；植物突变体：拟南芥的花形态呈现表观遗传相关的表型，两植物的基因组相同，表观遗传调控导致突变；(经允许引自 Jackson et al. 2002 [©Macmillan])；克隆猫：基因组完全相同，但是毛皮颜色的表型不同（经允许引自 Shin et al. 2002 [©Macmillan])。

卵双生的双胞胎有着一样的 DNA 序列，他们表型的相同性也常常被用来强调遗传学的决定力量。然而，双胞胎的外部表型仍然存在差异。这很有可能是由个体生活中累积的表观遗传修饰所造成（Fraga et al. 2005）。因此，我们对表观遗传学在决定细胞命运，个体特征和表型中的重要性的理解还远远不够。就组织再生和衰老而言，这些过程到底是由细胞遗传程序决定的还是由表观遗传修饰决定的还并不清楚。全球性的表观遗传学研究热潮充分表明这是后基因组时代的一个关键的新前沿。

有人说过“人绝不仅仅是基因的简单累加”（Klar 1998），还有人说“你可以继承

DNA 序列之外的一些东西，这正是现在遗传学中让我们激动的地方”（Watson 2003）。编著这本书的最大动机就是全体作者深信能将这一激动带给新一代的学生、科学家和医生们，因为他们中的大部分接受过遗传学教育，却还没有学过掌控染色体分离和遗传决定的原则——表观遗传学。

## 2. 研究表观遗传学的模式系统

表观遗传学的研究同样需要合适的模式生物，这些模式生物的特性乍看往往与人类或哺乳动物细胞相去甚远。然而，从很多这类模式系统中我们已经获得大量极有价值的信息。第 1、2 章的历史回顾就给我们介绍了几个里程碑性的发现，这些发现始于早期的细胞生物学、遗传学、分子生物学和染色体介导的基因调控研究开始取得进展的阶段。很多模式生物在深入挖掘和最终解决表观遗传学问题的过程中起着关键的作用。更值得一提的，在不同的模式生物中得到的看似相隔甚远的发现最终能够相互结合、贯通，为整个表观遗传的研究整体服务。这部分的目的就是重点描述其中的重大发现，细节的部分会在本书的后续章节中详细阐述。读者了解这些发现时，希望能关注研究者使用这些模式生物时所运用的基本概念，他们的贡献在于提出了常用的概念，而不是细节问题（图 3-2）。

“低等”真核生物酿酒酵母（*S. cerevisiae*，详见第 4 章）、裂殖酵母（*S. pombe*，详见第 6 章）以及链孢霉（*Neurospora crassa*）具有短暂的生活史，为遗传分析提供了强有力的工具。酿酒酵母和裂殖酵母生活史中发生的接合型转换（MAT switching）为染色体介导的基因调控的重要性提供了重要的验证实例。酿酒酵母出芽过程中，特有的沉默信息调控因子被证明与有特定修饰的组蛋白相互作用，而组蛋白先前已被证实参与基因调控（Clark-Adams et al. 1988；Kayne et al. 1988），在裂殖酵母的裂殖过程中，通过组蛋白修饰来激活或抑制信号的模式与多细胞动物非常相似。这为寻找抑制或增强基因沉默的蛋白提供了强大的遗传筛选途径。最近，在裂殖酵母研究中提出了一种机理，该机理将组蛋白修饰的诱导与 RNAi 联系起来（Hall et al. 2002；Volpe et al. 2002）。很短的时间内，RNAi 机制被发现参与拟南芥的转录水平的基因沉默调控，预示着这种调控机制对很多生物体调控都有重要意义。

其他非经典的生物体也对解答未知的表观遗传途径有着或多或少的贡献，一开始人们认为这种途径是很特殊的。真菌类的链孢霉显现出的重复序列诱导的点突变（RIP）是比较特殊的非孟德尔现象，也成为研究表观遗传控制的一个模型（见第 6 章）。后来这个生物被用来证明了组蛋白修饰和 DNA 甲基化之间的第一个功能上的联系（Tamaru and Selker 2001），这也是后来被扩展到“更高等”生物体的一个发现（Jackson et al. 2002）。纤毛类原生动物，如四膜虫和草履虫，常常被生物学实验室用作方便的显微镜标本，而且加速了重要的表观遗传发现，这是因为它们具有独一无二的核二形性。每个细胞有两个细胞核，一个转录活跃的体细胞大核和一个转录不活跃的生殖细胞微核。大核可以作为“活跃”染色质的充分的起始材料，从中纯化到了首个细胞核组蛋白修饰酶——一种组蛋白乙酰基转移酶（HAT）（Brownell et al. 1996）。纤毛生物还有一种

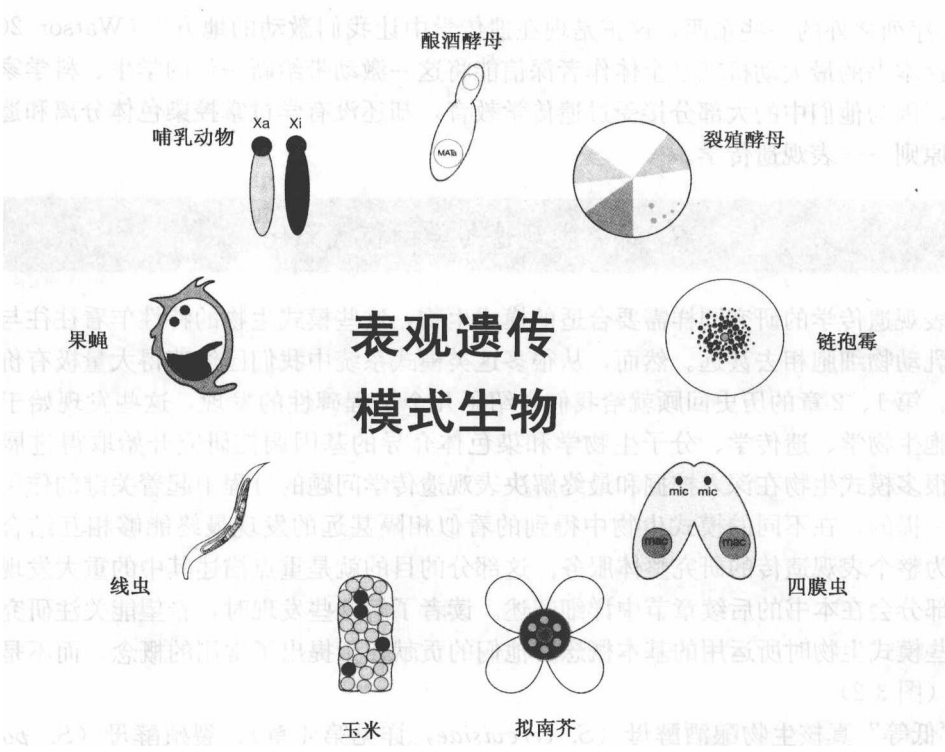



图 3-2 表观遗传学研究中用到的模式生物

酿酒酵母：接合型转变用于研究表观遗传的染色体调控；裂殖酵母：花斑基因沉默可以表现在菌落的聚集上；链孢霉：表观遗传的基因组防御系统包括重复序列诱导的点突变，平抑作用，以及减数分裂中对于不配对 DNA 的沉默，这反映出 RNA 干扰途径、DNA 和组蛋白甲基化之间的相互联系；四膜虫：体细胞核生殖系细胞核中染色体可由表观遗传学调控的机制区分；拟南芥：用于研究 DNA、组蛋白和 RNA 介导的阻遏机制引起的抑制作用的模型；玉米：用于研究印记、副突变和转座子引起的基因沉默的模型；线虫：生殖细胞系中表观遗传调控；果蝇：由白眼基因区块化表达或沉默显现出位置效应花斑；哺乳动物：X 染色体失活。

闻名的特殊现象，就是它们的有性生殖周期中会有小的非编码 RNA 和组蛋白修饰引起的程序化 DNA 剔除（见第 7 章）。

在多细胞生物中，从无脊椎类（线虫、果蝇）或植物（拟南芥）到“高等一些”的生物，再到一些“更高等”的脊椎动物（哺乳类），基因组大小和生物体复杂度一般来说都在增加。然而植物却成了表观遗传领域的楷模，提供了尤其多的表观遗传发现的资源（第 9 章）：由转座元件、副突变（McClintock 1951）到首次报道的转录沉默有关的非编码 RNA（Ratcliff et al. 1997）. 从植物的研究中还得到了 DNA 甲基化、组蛋白修饰以及 RNA 干扰系统之间的重要联系。人们发现了植物表观等位基因，它们有着滑稽的名字如“超人”（*SUPERMAN*）、*KRYPTONITE*（Jackson et al. 2002）和一些春化基因（Bastow et al. 2004；Sung and Amasino 2004），这进一步加深了人们对表观遗传学和细胞记忆在发育上的作用的理解和见地。植物茎尖细胞也使得人们有机会研究一些

如体细胞重生以及干细胞可塑性之类的重要问题（见第9、11章）。

对于理解动物的发育，果蝇是一个很早就有并且一直被利用的遗传“生产线”。在Muller（1930）先驱工作的基础上，人们获得了许多发育突变体，包括以下要讨论的同源异型的转化以及位置效应花斑（PEV）的突变体（见第5章）. 同源异型转化突变体使人们认为可能有建立和维持细胞身份/记忆的调控机制，后面发现这是由Polycomb和Trithorax系调控的（见第11、12章）。对于PEV，基因活性是由周围的染色质结构控制的，而不是一级DNA序列。这个系统是研究表观遗传调控因子的一个重要资源库（见第5章）。人们认为有超过100个花斑抑制［Su（Var）］基因编码异染色质的组分。没有这些里程碑性的研究基础，也就不可能发现第一个组蛋白赖氨酸的甲基转移酶（HKMT）（Rea et al. 2000）和后来在组蛋白赖氨酸甲基化研究中得到的进步。按照生物学中常有的做法，人们也相应地筛查了酿酒酵母和植物，并发现了与果蝇*Su*（*var*）功能保守的沉默突变体。

在线虫中用RNA干扰文库的反向遗传学使我们进一步理解了后生动物发育中的表观遗传调控。其中，有细化了的缺失细胞所有发育途径的追踪细胞命运的综合研究，这突出了Polycomb和Trithorax系统很可能伴随着多细胞化的出现而产生（见本章第12、13节）。这些表观遗传调控机制对生殖细胞中的基因调控尤其有着至关重要的作用（见第15章）。

尽管一些表观遗传学研究是在人类细胞系和原代细胞培养物中开展的，但表观遗传在哺乳类发育过程中的作用大多数是在小鼠中阐明的。基因“敲除”和“敲入”技术的来到为关键表观遗传调控子的功能性分析提供了工具。例如，Dnmt1 DNA甲基转移酶突变小鼠为哺乳类DNA甲基化的作用提供了功能上的参考（Li et al. 1992）。这种小鼠是胚胎期致死的，并且显示了受损的印记形式（见第18章）。破坏DNA甲基化也显示可以导致基因组不稳定以及转座子活性的重新活化，尤其是在生殖细胞中（Walsh et al. 1998；Bourc and Bestor 2004）。大约有100个已确定的染色质调控因子（如组蛋白和DNA修饰酶类，核小体重塑复合体以及RNA干扰系统的组成成分）被在小鼠中敲除，产生的突变体表型影响到细胞增殖、属系化、干细胞可塑性、基因组稳定性、DNA修复以及染色体分离进程，这些在体细胞和生殖细胞系中都有出现。因此大多数突变体也在疾病发展和癌症中涉及也并不奇怪了。所以，许多表观遗传调控中关键性进展，至少利用了以上提到的许多模式生物的独特生物学特征。没有以上模式生物的这些生物学进展和功能性分析（遗传学和生物化学的），也就不可能有今天表观遗传调控研究获得的许多进展。

## 3. 表观遗传学的界定

以上讨论引出了这样一个问题：是什么共同的线索将多样的真核生物从表观遗传调控的角度联系到一起的呢？事实上，不同的表观遗传现象大部分是通过真核生物中的DNA不是“裸露的”这一事实联系在一起的。DNA与特异的蛋白质一起构成了染色质。染色质的最简单形式——组蛋白与缠绕的DNA组成的核小体单位（Kornberg

1974)，起初被认为是一种包装和组织 DNA 的被动包装分子。后来发现了染色质的不同形式，这是通过共价和非共价机制相结合而产生的，这些机制很快地被揭示了（见第 6 章）。这包括许多翻译后的组蛋白修饰，移动或改变核小体结构的能量依赖的染色质重塑步骤，新的组蛋白变体动态进出核小体以及小的非编码 RNA 的靶向作用。在许多高等的真核生物中，DNA 自身也可以被共价修饰，一般通过胞嘧啶的甲基化来实现，一般为 CpG，但并不总是如此。总的来说，这些机制提供了一套相互作用的途径，它们都引起了染色质多聚体的变化（图 3-3）。

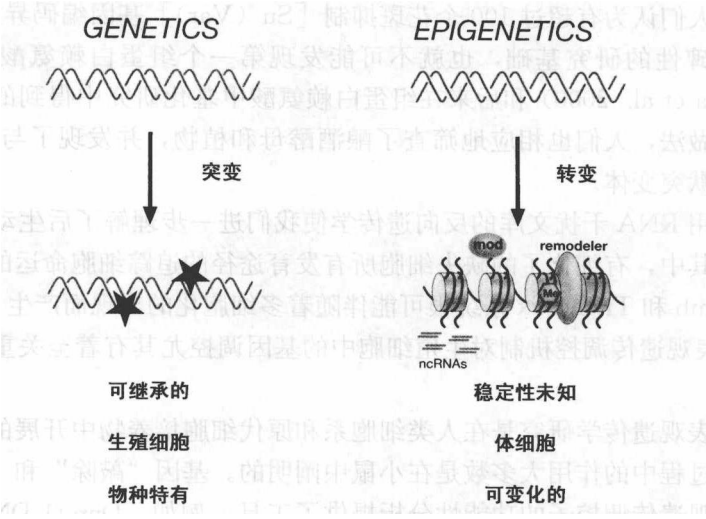



图 3-3 遗传学与表观遗传学

遗传学：DNA 模板（绿色螺旋）的突变（星号）能够通过体细胞和种系遗传。表观遗传学：染色质结构的变化通过以下方式调节基因组的功能：①组蛋白修饰（mod）；②染色质重塑（remodeler）；③组蛋白组分改变（黄色标示核小体）；④DNA 甲基化（Me）；⑤非编码 RNA。染色质上的这些标记在细胞分裂过程中能够被继承并共同作用决定细胞表型。

这些修饰及染色质变化有许多是可逆的。因此它们不太可能通过生殖细胞被继承。由于它们可以应对内外刺激使染色质模板产生变化，过渡性的表观遗传标记物似乎很有吸引力（Jaenisch and Bird 2003），因为它们可以调节转录机器的进入和/转录延伸，这些是转录机器“阅读”DNA 模板所必需的条件（Sims et al. 2004；见第 10 章）。一些组蛋白修饰（如赖氨酸甲基化）、DNA 甲基化区域和改变了的核小体结构却可以在几代细胞分裂中都保持稳定。这些修饰建立了“表观遗传状态”或者细胞记忆的方式，但我们对这些仍然了解得太少。从这一点来说，染色质“标志”可以被看做是一个高度组织有序的信息存储系统，能够索引到基因组的不同区域而且还可以应答环境信号并控制基因的表达程序。

染色质模板存在的重要意义在于它提供了解读 DNA 信息的多个维度和层面，并且能够加强遗传信息。考虑到真核生物基因组的庞大和复杂程度，特别是对于多细胞生物体，这可能是一个必要条件（第 11 节有更详细的论述）。在这样的生命体中，单个受精卵在发育过程中不断地发展，从一个受表观遗传调控的单基因组发展成存在于 200 多种

不同类型细胞中的多种表观基因组（图 3-4）。这种程序性的变化被认为组成了一种“表观遗传密码”，这一密码使经典的遗传密码中所隐藏的信息产生了意义非凡的扩展（Strahl and Allis 2000；Turner 2000；Jenuwein and Allis 2001）。虽然这是一个引人注目的假说，但需要强调的是，我们还需要做更多的工作去验证它以及一些相关的有争议的学说。其他一些观点也有所发展，这些观点认为已经具有一定系统的“组合密码”，如三联体密码，并不存在于组蛋白中或还远未认识到这一现象的存在（Schreiber and Bernstein 2002；Henikoff 2005）。除去这些不确定的因素外，我们均倾向于一种被普遍接受的观点，即，一种共价和非共价的机制相互结合促使产生了一种染色质状态，使其在细胞的分化和发育过程中能够作为模板，但其中涉及的机制才刚刚被人们加以阐释。但这些不同的染色质结构在 DNA 复制和细胞有丝分裂过程中是如何准确地传递下去仍然是一个有待于解决的根本问题。

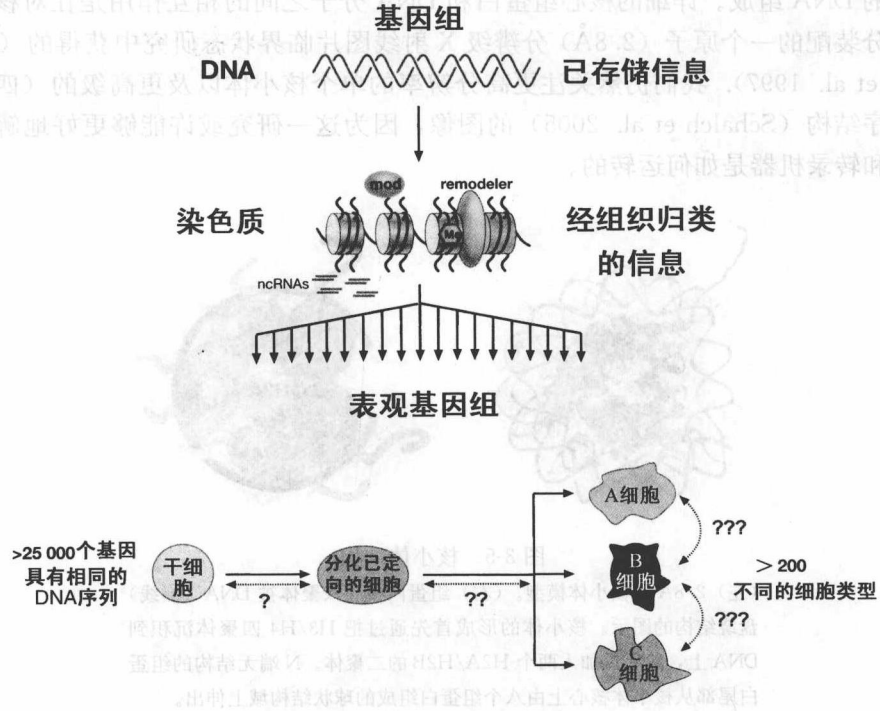



图 3-4　DNA 和染色质

基因组：某一个体不变的 DNA 序列（双螺旋）。表观基因组：染色质的总体构成，对应任何特定细胞中的整个染色体组。它会随细胞类型的不同而变化，以及对其收到的内外界信号发生反应。表观基因组会在多细胞生物由一个干细胞（受精卵）发育到许多已分化细胞这一过程中发生变化。分化或去分化的转变需要细胞的表观基因组重编程。

多细胞生物体在发育过程中细胞与细胞之间的表型差异被 Waddington 描述成是“表观遗传学全貌”（Waddington 1957）。在从干细胞到完全分化的众多细胞尽管它们都具有一致的 DNA 序列，但是它们的基因表达却相差悬殊。正是基于这些知识，表观遗传学后来被定义为“不依赖 DNA 序列差异的核继承性”（Holliday 1994）。

自 DNA 双螺旋的发现和早期对表观遗传学的解释以来，我们对表观遗传控制和它潜在机制的理解取得了巨大的进步，以致有人开始用更高级的词汇如“领域”而不是“现象”来描述它（Wolffe and Matzke 1999；Roloff and Nuber 2005；第 1 章）。在过去的十年中，有关修饰染色质的众多酶家族的研究取得了相当大的进展。因此，在当今的词汇中，表观遗传学能够在分子（或机理）角度定义为“在同一基因组上建立并将不同基因表达（转录）模式和基因沉默传递下去的染色质模板变化的总和。”

## 4. 染色质模板

核小体是基本的染色质重复单位（Kornberg 1974）。一方面，这一染色质基本单位由核心组蛋白（H2A、H2B、H3 和 H4）组成的一个蛋白八聚体和一段 147 bp 包绕在其外周的 DNA 组成。详细的核心组蛋白和 DNA 分子之间的相互作用是在对核小体在重组部分装配的一个原子（2.8Å）分辨级 X 射线图片临界状态研究中获得的（图 3-5）（Luger et al. 1997）. 我们仍然关注更高分辨率的单个核小体以及更高级的（四聚核小体）有序结构（Schalch et al. 2005）的图像，因为这一研究或许能够更好地解释染色质重塑和转录机器是如何运转的。

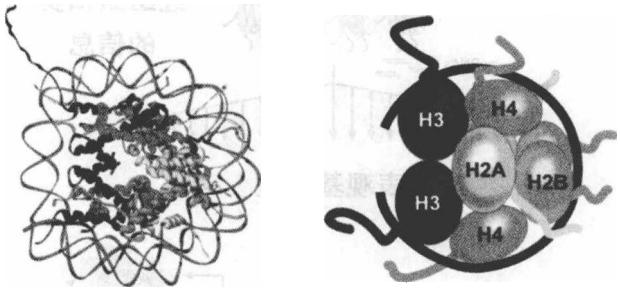



图 3-5 核小体结构

（左）2.8Å 的核小体模型。（右）组蛋白核心八聚体被 DNA（黑线）盘绕结构的图示。核小体的形成首先通过把 H3/H4 四聚体沉积到 DNA 上，然后再加入两个 H2A/H2B 的二聚体。N 端无结构的组蛋白尾部从核小体核心上由八个组蛋白组成的球状结构域上伸出。

组成核小体的核心组蛋白是一类小分子强碱性蛋白。它们均有球状结构域和可变的（相对无结构的）从核小体表面伸出的“组蛋白尾部”组成（图 3-5）。从酵母到人，组蛋白的氨基酸序列是非常保守的。这种高度的保守性也支持一些基本的观点：这些组蛋白，即使在无结构的尾部区域也可能具有关键的功能。事实上，这些尾部区域，尤其是组蛋白 H3 和 H4 的尾部区域，为研究核小体（乃至染色体）的可变性提供重要的线索，如一些残基受到广泛的翻译后修饰（请见本书附录 B 中一些已知的组蛋白修饰和规范的命名方法）。

核心组蛋白的乙酰化和甲基化，主要是 H3 和 H4，是应该首先描述的共价修饰，并一直被认为与转录活性的正负改变相关。自 Allfrey 及其合作者最先开始（Allfrey et

al. 1964)，很多类型的共价修饰已经被鉴定和描述出来。其中包括组蛋白的磷酸化、泛素化、sumo 化（sumoylation）、ADP 核糖基化（ADP-ribosylation）、生物素化、脯氨酸异构化（proline isomerization），以及其他可能的将要被描述的修饰（Vaquero et al. 2003）。这些修饰发生在特异的位点和残基上，其中一些在图 3-6 和附录 B 中列出。特定的酶和酶的复合物（其中一些在接下来的综述和单个章节中被重点描述）催化这些共价标记。这些列表将在接下来的几年中继续增加，因此我们只描述那些能勾画出我们认为重要的一般概念和原则的标记和酶。

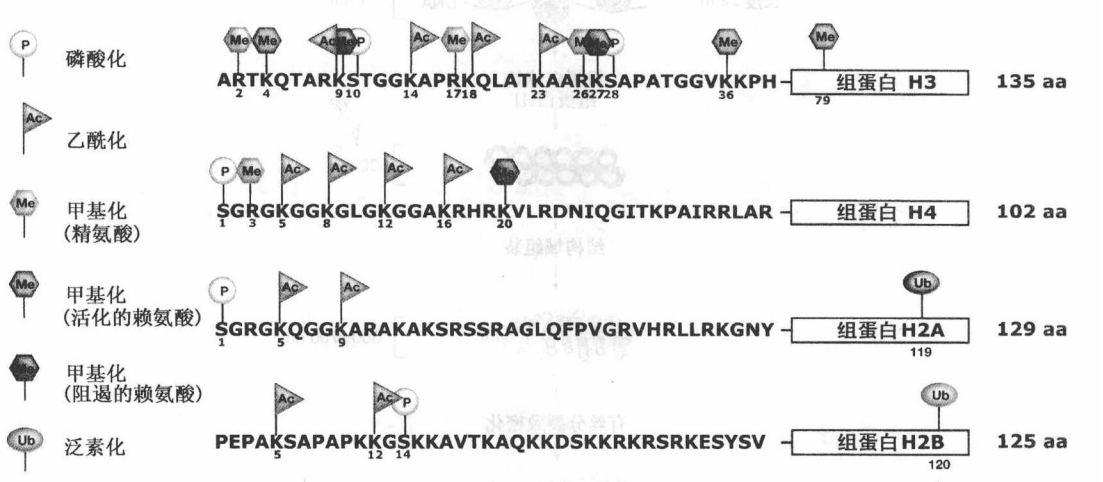



图 3-6　组蛋白尾巴的修饰位点

组蛋白的 N 端尾巴是核小体堆栈的四分之一，它们囊括了已知共价修饰位点的大部分。修饰也发生在球状结构域（矩形图示），其中一部分见图标注。一般来讲，活化标签包括乙酰化（青色 Ac 三角），精氨酸甲基化（黄色 Me 六角）以及一些赖氨酸甲基化，如 H3K4 和 H3K36（绿色 Me 六角）。球状区域的 H3K79 具有抑制沉默的功能。抑制标签包括 H3K9、H3K27 和 H4K20（红色 Me 六角）。绿色：活化标签，红色：抑制标签。

在某些染色质区域，核小体可能包含代替核心组蛋白的组蛋白变体，现行研究表明这种组成的差异有助于染色体的标记区域具有特定功能。

## 5. 染色质的高级结构

DNA 和核小体的聚合体——染色质，是一种包含多种结构的动态分子。染色质包括常染色质和异染色质，细胞学家在用 DNA 染料进行染色时，观察到两种不同的染色模式，据此将染色质分为常染色质和异染色质两种类型。常染色质是松散的染色质，它既可以是转录活跃状态，也可以是转录非活跃状态。异染色质通常被认为是高度致密和沉默的染色质，它有两种形式：组成型异染色质，它是永久沉默的染色质，基因在生物体的所有细胞类型中都几乎不表达；兼性异染色质，在某些细胞的特殊细胞周期或发育阶段其基因表达会被抑制。大量的染色质状态和大量的文献暗示了染色质是一种高度动态的大分子结构，当它接收到上游相关的生理信号时会倾向于重塑或重建。直到最近，关于支配重塑过程的分子机制才有了较大的进展。

课本中提到的 11nm 串珠状结构代表一种活跃的极大程度上松散的间期染色质构造，在这种结构中，DNA 周期性的缠绕在核小体周围（图 3-7）。然而染色质纤维并不总是由规则间隔的核小体排列而成。在原子分辨率下，人们开始观察到核小体可以被不规则地包装和折叠成高级结构（Khorasanizadeh 2004）。在细胞命运决定或细胞周期的不同阶段（间期和有丝分裂期），有差异的和高级的结构会在基因组的不同区域形成。

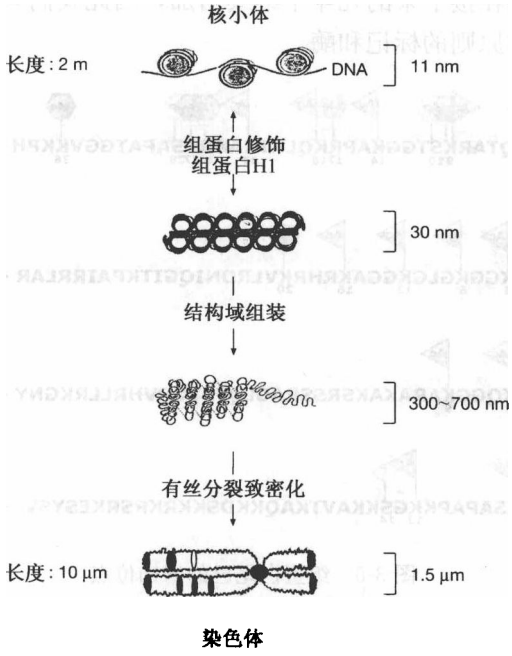



图 3-7 染色质的高级结构

11nm 纤维代表 DNA 缠绕在核小体上形成的结构。30nm 纤维是在组蛋白 H1 存在下，进一步包装成的尚有待证实的结构（本图示以其螺线管模型）。300～700nm 纤维代表在间期和中期染色质上存在的动态的高级环状结构。1.5μm 致密的染色体代表仅在核分裂（有丝分裂或减数分裂）时才出现的最紧密的染色质结构。目前还不清楚有丝分裂的染色体带型（如 G 带型和 R 带型）与哪种染色质结构相对应。

在 11nm 的模板上核小体的排列会被共价修饰了的组蛋白尾巴的顺式或反式效应所改变（图 3-8）。顺式效应是由被修饰的组蛋白尾巴的物理特征的改变引起的，例如静电电荷的调整，反过来尾巴结构也可以改变核小体间的联系。例如组蛋白乙酰化，长时间以来被认为能够中和高碱性组蛋白尾巴的正电荷，这使染色质纤维产生局部扩张，从而使转录机器更容易接近 DNA 双螺旋。磷酸化通过净负电荷的增加形成“电势阱”（Dou and Gorovsky 2000），从而通过改变染色质聚合体高级折叠状态来改变核小体包装或暴露组蛋白氨基末端。同样，连接组蛋白 H1 被认为通过屏蔽相邻核小体之间的 DNA 连接中的负电荷来促进高级结构纤维的包装。相对较大的化学修饰，例如泛素和 ADP 核糖的参与也会诱导组蛋白尾巴的不同组装并打开核小体排列。组蛋白尾巴能在多大程度上通过修饰依赖性和修饰非依赖性的机制诱导染色质包装仍不清楚。

组蛋白修饰可以通过招募一些结合这些特定修饰的蛋白质到染色质，产生我们所说

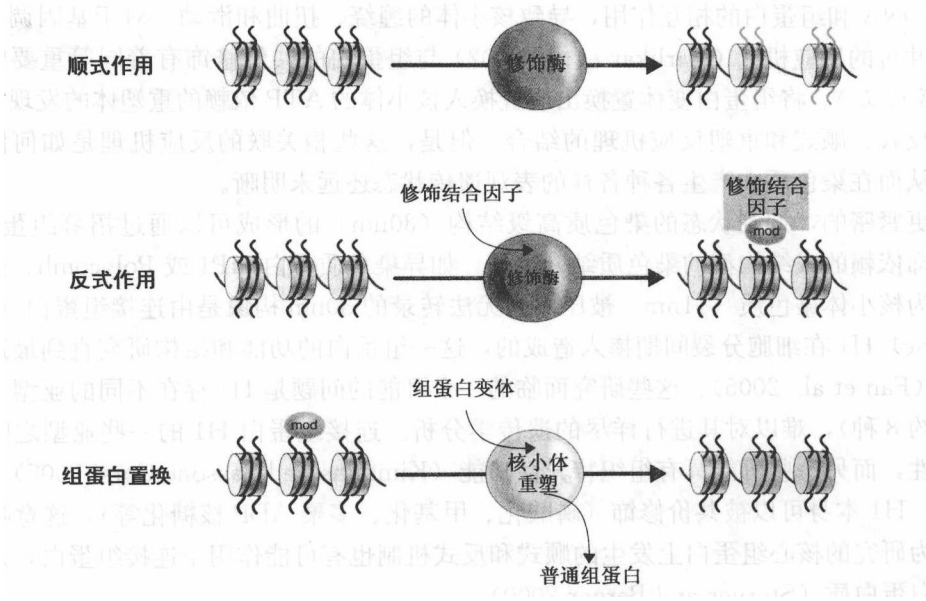



图 3-8　染色质模板的转换（顺式/反式）

顺式作用：组蛋白尾巴上的共价修饰引起结构或电荷的改变，造成染色质组织结构的变化。反式作用：组蛋白尾巴上的酶促修饰（如 H3K9 甲基化），导致染色质结合蛋白（修饰结合物，如 HP1）的亲和性结合。修饰结合物（或其所在的蛋白复合物）的结合，引发染色质结构的进一步改变。组蛋白置换：组蛋白共价修饰可以通过核小体重塑交换复合物（nucleosome-remodeling exchanger complex）实现组蛋白变体对某个核心组蛋白的置换。

的反式效应。这可以理解为，在特定的背景下，“读取”某个组蛋白共价修饰所蕴藏的信息。某些结合蛋白具有专一的亲和力，可以结合在特定的组蛋白尾巴。它就像粘扣一样粘在特定的修饰上，还往往会带来一个更大的蛋白质复合体，使其能作用于染色质多聚体。例如，溴区结构域（bromodomain），一段可以识别组蛋白乙酰化残基的序列，常常存在于染色体重塑蛋白复合体的乙酰基转移酶（HAT）中（见第 7 节，图 3-10 ）(Dhalluin et al. 1999；Jacobson et al. 2000)。同样，组蛋白尾巴上甲基化的赖氨酸可以与 chromo 结构域（chromo domain）（Bannister et al. 2001；Lachner et al. 2001；Nakayama et al. 2001）或其他类似的结构域结合，如 MBT 或 tudor（Maurer-Stroh et al. 2003；Kim et al. 2006），以促进下游的染色质调节事件。某些情况下，含有 chromo 结构域的蛋白可以通过结合组蛋白甲基转移酶（HKMT）来甲基化相邻的核小体，而甲基化的核小体又能再次被 chromo 结构域所识别，从而导致异染色质的蔓延（见第 5 章）。

组蛋白在尾巴或核心区域的修饰（Cosgrove et al. 2004）都可以招募 ATP 依赖的染色质重塑复合体到 11nm 染色质纤维，促进从可转录的常染色体状态到转录活跃状态的改变。这种转变可以通过几种方式实现，例如，组蛋白八聚体在 DNA 上的滑行，DNA 环化导致其在组蛋白上缠绕方式的转变（见第 12 章），或以组蛋白变体置换某些核心组蛋白（见第 13 章）。ATP 依赖的染色体重塑体（如 SWI/SNF）把水解释放的能量用于

改变 DNA 和组蛋白的相互作用，导致核小体的缠绕、扭曲和滑动。对于基因调节，这些非共价的反应机制（Narlikar et al. 2002）与组蛋白的共价修饰有着同等重要的作用（见第 10 章）。将组蛋白变体置换出或置换入核小体的 ATP 依赖的重塑体的发现，实现了将反式、顺式和重塑反应机理的结合。但是，这些相关联的反应机理是如何协同作用，从而在染色质中产生各种各样的表观遗传状态还远未明晰。

更紧密的，抑制状态的染色质高级结构（30nm）的形成可以通过招募组蛋白 H1 和修饰依赖的或结构型的染色质结合因子，如异染色质蛋白 HP1 或 Polycomb。虽然普遍认为核小体染色质（11nm）被压缩到无法转录的 30nm 构象是由连接组蛋白（linker histone）H1 在细胞分裂间期掺入造成的，这一组蛋白的功能和结构研究直到最近都很困难（Fan et al. 2005）。这些研究面临的一个可能的问题是 H1 存在不同的亚型（哺乳类中约 8 种），难以对其进行详尽的遗传学分析。连接组蛋白 H1 的一些亚型之间存在冗余性，而另一些可能具有组织特异性功能（Kimmins and Sassone-Corsi 2005）。有趣的是，H1 本身可以被共价修饰（磷酸化、甲基化、多聚 ADP 核糖化等），这意味着目前广为研究的核心组蛋白上发生的顺式和反式机制也有可能作用于连接组蛋白，乃至非组蛋白蛋白质（Sterner and Berger 2000）。

关于 30nm 染色质的具体组织方式已经有许多辩论。大体上，“螺线管”（单起始螺旋）模型，即核小体绕中轴缠绕（6～8 个核小体/圈），或者更开放的“之字形”模型——采用较高阶的自装配方式（双起始螺旋），都已经被描述过。新的证据，包括从含四个核小体的模型获得的 X 射线结构，倾向于由双起始的之字形连接 DNA 将两层核小体相联结（Khorasanizadeh 2004；Schalch et al. 2005）。尽管有这样的进展，我们注意到连接组蛋白不在这一结构中，而且即使它存在，30nm 染色质结构只将 DNA 压缩了约 50 倍。因此，决定中期或分裂期染色质状态的众多更高级的染色质组织方式尚有待于被光学和电子显微镜以外的手段来解决。尽管结构上尚不确定，最近在活细胞中的成果已经确立了间期染色质 30nm 以上多种折叠方式的存在。一个值得关注的进步是标记活细胞内特定 DNA 序列的方法学发展，使得实时观察染色质打开和关合的动态过程成为可能。有趣的是，这些结果揭示了通过染色质重塑正负调控因子的动态互作，可以改变染色质高级结构使之有利于或不利于基因表达（Fisher and Merkenschlager 2002；Felsenfeld and Groudine 2003；Misteli 2004）。

更大的环状染色质结构（300～700nm）的组装，可能是通过将染色质纤维经染色质结合蛋白例如核纤层蛋白（nuclear lamin）锚定到核周或者其他核骨架。这些结构是否能产生功能上有意义的“染色体界域”（chromosome territories）尚不清楚，但是许多报道表明这一概念值得认真关注。例如，观察到了转录活跃的多个染色质位点被聚集到 RNA 聚合酶 II 转录因子周边的现象，而且相似的概念似乎也可被用于解释在复制活跃的 DNA 和 DNA 聚合酶周边的聚集现象。此外，也有报道表明，与此相对的，沉默的异染色质（尤其是着丝粒周边区域）以及位于反式区域的基因也有类似的聚集现象（见第 4、12 章）。这一现象是如何被控制的以及染色质区域的核定位对基因组究竟有何调节作用仍不清楚。尽管如此，越来越多的证据表明染色质的激活或沉默与其在核内所处的特定位置有关（Cremer and Cremer 2001；Gilbert et al. 2004；Janicki et al. 2004；

Chakalova et al. 2005)。

结构最致密的DNA出现在有丝分裂和减数分裂中期。以确保每个子代细胞可以以染色体的形式得到正确拷贝数的基因组（一个或两个拷贝，取决于细胞分裂的种类)。这种致密化过程可以惊人地将完全伸展时长约2m的DNA压缩到若干直径为1.5μm的染色体（图3-7)。这一约10 000倍以上的压缩过程涉及组蛋白H1、H3的超磷酸化，ATP依赖的致密因子（condensin）复合物，黏结蛋白（cohesin）复合物，以及拓扑异构酶II。非组蛋白复合物是如何作用于分裂期染色体（分裂期染色质修饰）以及在细胞周期内这些复合物是怎样与染色体结合及分离的具体机制仍然需要进一步确定（Bernard et al. 2001；Watanabe et al. 2001)。众所周知的有丝分裂期组蛋白H3和H1家族蛋白的磷酸化（如Ser10和Ser28）可能为此提供了很好的线索，但是这些有丝分裂期的标记的生物学作用尚有待于被遗传学和生物化学实验确定。有趣的是，“双元开关”理论提出组蛋白上特定的甲基化标记与更动态可逆的磷酸化标记可以作为一种二元的开关共同调节下游修饰结合蛋白与染色质的结合和释放（Fischle et al. 2003a)。HP1与组蛋白H3甲基化的第九位赖氨酸的结合受Ser10的磷酸化作用调节，这一范例支持“甲基化/磷酸化开关”的假说（Daujat et al. 2005；Fischle et al. 2005；Hirota et al. 2005)。

特殊的染色体区域，如端粒和着丝粒，执行调节染色体动态过程的特定功能。端粒作为染色体的末端，保护着DNA分子并提供了最末端DNA的复制途径。在细胞核分裂时，着丝粒为纺锤体微管蛋白提供结合位点。这两种特殊区域对于染色体分离的精确性都起着至关重要的作用。有趣的是，端粒和着丝粒异染色质都与常染色质，乃至其他异染色质有着显著的区别，因为它们有着特殊的染色质结构抑制基因表达及DNA重组。当常染色质基因被人为插入到着丝粒或者端粒区，会造成这些基因的沉默。这为筛选先前描述过的位置效应花斑现象（PEV）以及端粒位置效应（TPE；Gottschling et al. 1990；Aparicio et al. 1991）的抑制基因和增强因子提供了强有力的工具。着丝粒和端粒有着特定的分子特征，如它们组蛋白乙酰化水平很低。有趣的是，着丝粒还有另外一个标志——组蛋白变体CENP-A。组蛋白变体CENP-A在染色体分离中起到重要作用。因此，正确的组装和维持特征性的着丝粒和着丝粒周边异染色质对有丝分裂和减数分裂的完成，以及细胞的生存至关重要。在已经研究得比较透彻的着丝粒和着丝粒周边组成型异染色质的基础上，关于着丝粒和端粒形成的表观遗传学调控机制研究也取得了进展。一些精妙的实验表明新着丝粒（neocentromere）可以替换正常的着丝粒，表明DNA序列并非着丝粒特征性的决定因素。相反，着丝粒这种特殊的染色体区域是由表观遗传标记，包括着丝粒特征性修饰和组蛋白变体所决定。染色质编码区非编码区和重复区域对表观遗传学标记的调控机制研究取得了相当的进展。这些机制是否或怎样与染色体带型相关尚不清楚，但仍然给我们一种诱人的可能性。对于这些独特的染色体区域的表观遗传调控机理研究是非常必要的，因为大量事实表明许多癌症都与基因组不稳定性有关，这种基因组不稳定性是某些疾病恶化和肿瘤形成的显著特征。

## 6. 常染色质与异染色质的区别

我们知道常染色质和异染色质各自都有许多存在形式，以下将分别讨论。常染色质（或称“活性”染色质）主要由只占了哺乳动物基因组的一小部分（少于4%）的编码序列构成。具有转录潜能的编码序列被哪些分子信号所标识？染色质结构在这一过程中又有何作用呢？大量的文献表明，常染色质是以一种开放的（非浓缩的）、对核酸酶更为敏感的构象存在，这样它就能随时开启基因表达，即便有时它仍存在于非转录状态。有些基因的表达是无所不在的（看家基因）；另外一些基因则是受发育的调控或是受胁迫诱导以便对环境信号进行应答。顺式作用DNA元件（启动子、增强子和基因座控制区）和反式作用因子相互协作，与RNA聚合酶和各种相关因子一起促进基因的转录（Sims et al. 2004）。这些作用因子都在进化过程中经历了严格的选择，以保障那些繁多复杂的生物化学反应发生在正确的空间和时间。是否染色质提供了一个“索引系统”来更好地保证这些转录机器能够在适当的细胞类型中正确地接近它们的目标序列呢？

在DNA水平上，富含AT的启动子区通常缺少核小体，可能并非严格地以经典的B型DNA的形式存在，以利于转录因子的结合（Mito et al. 2005；Sekinger et al. 2005）。然而，转录因子与DNA的结合并不足以启动转录。目前已知的研究表明，通过诱导产生与转录活性相关的组蛋白修饰（如乙酰化和H3K4甲基化），进而招募核小体重塑因子，能够促进转录机器的参与（图3-9，第10章）。当转录机器解旋并转录通过染色质纤维后，组蛋白会被置换为组蛋白变体，以保证染色质模板的完整性（Ahmad and Henikoff 2002）。然而要想得到成熟的mRNA，还需要经过转录后修饰，包括剪接、加多聚腺苷酸和出核转运。因此，常染色质的综合定义或许是指一个复杂的染色质状态，包括了各种动态的和精细的机器，它们既相互作用也与染色质纤维紧密作用，共同转录出有功能的RNA产物。研究转录活化机器（activating machinery）与转录设备以及染色质模板间相互作用的一般规律是当前一个令人振奋的研究领域。尽管基于动态性特征，它严格来讲不能被归类于表观遗传学，而更多的是关于转录和染色质动态学的研究。

那么异染色质又该如何定义呢？虽然历史上对异染色质的研究明显少于常染色体，但是最新的研究表明从酵母到人类（虽然酿酒酵母有不同的异染色质形态），异染色质在基因组的组织和机能起着至关重要的作用。96%的哺乳动物基因组是由非编码序列和重复序列组成，这一现象凸现了异染色质的重要性。关于异染色质形成机制的新研究揭示了意想不到的发现。例如，能够产生双链RNA的非序列特异性转录（non-sequence-specific transcription）会被一个类似于RNA干扰的机制所沉默（见第10节）。这些双链RNA的产生提供了一个警报信号，表明这些潜在的DNA序列不能产生有功能的产物，或者已经被RNA转座子或病毒侵袭。这些双链RNA被Dicer剪切后会被特定的复合体定位至染色质，继而通过一系列事件最终导致异染色质的形成。通过运用多种模式系统，科学家们在研究异染色质“禁锢”（locked-down）状态的形成通路方面取得了显著的进展，而这一通路在各种模式生物中是高度保守的。这条通路主要包括组蛋白尾

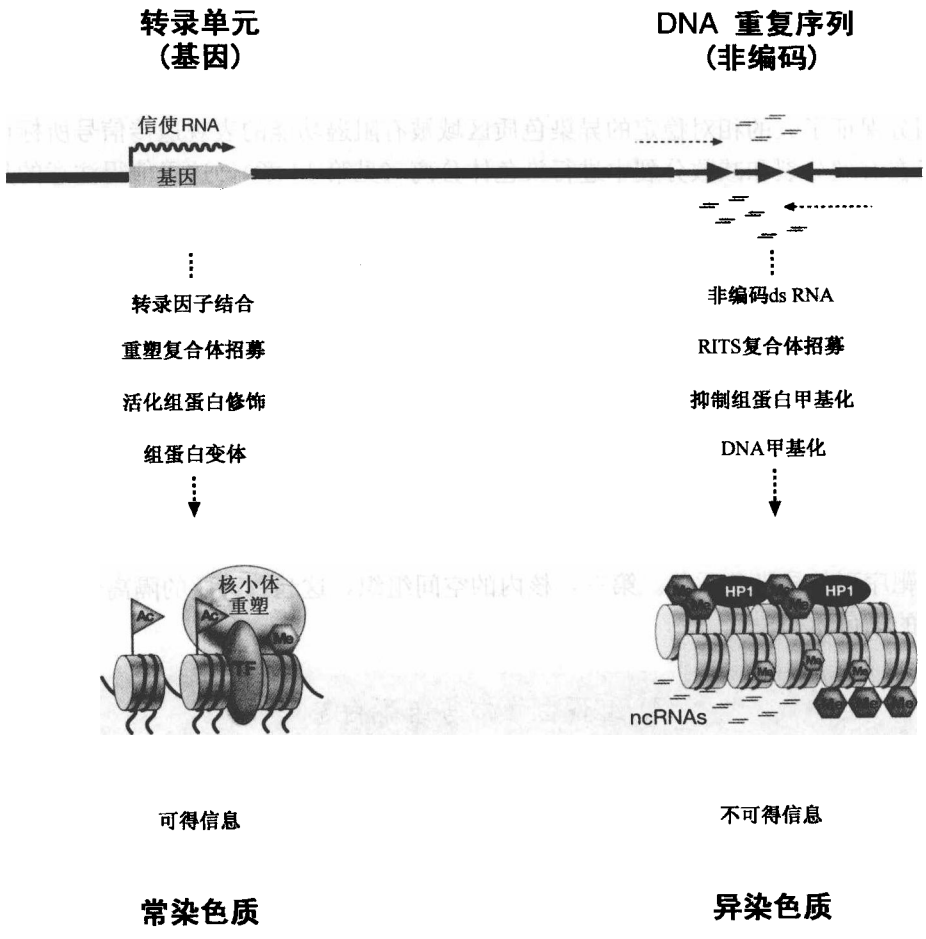



图 3-9　常染色质与异染色质的差别

常染色质和组成型异染色质之间一般差别的概述。包括转录产物的类型、DNA 结合蛋白（如转录因子）的招募、染色质结合蛋白和复合体、组蛋白共价修饰以及所包含的组蛋白变体等方面的区别。

巴的去乙酰化，某些赖氨酸残基的甲基化（如 H3K9），异染色质结合蛋白的招募（如 HP1）和 DNA 甲基化（图 3-9）的建立，虽然具体的顺序和细节可能会有所变化。似乎将某些基因组序列置于核内特定的沉默区域内可能有助于异染色质的形成。有趣的是，越来越多的证据表明，至少是在高等生物中，异染色质可能是一种默认状态。而强启动子或是增强子的存在能突破异染色质的束缚而产生转录产物。即使在低等生物中，异染色质形成的一般概念似乎也适用。它的特点包括组蛋白尾的低乙酰化，以及对乙酰化敏感的异染色质蛋白的结合（如 SIR 蛋白，详见第 4 章）。不同种类的真菌中（如芽殖和裂殖酵母），存在有不同程度的组蛋白甲基化和类似于 HP1 的蛋白。这些物种的基因组大部分处于默认的可转录状态，但仍有一些类似于异染色质的基因组区域（交配

区、端粒、中心粒等）存在，插入到这些区域的外源测试基因的转录和遗传重组会受到抑制。

异染色质有什么功能呢？着丝粒是一个组成型异染色质的区域，它与可遗传的染色质状态紧密相关，在进化上被认为是由染色体上重复序列最大的聚集过程所演变而来。这种划分保证了大的相对稳定的异染色质区域被有阻遏功能的表观遗传信号所标记，这有助于在有丝分裂和减数分裂中进行染色体分离（见第 14 章）。这里值得注意的是，一些生物体，如在裂殖酵母中，这些着丝粒重复序列和相应的表观遗传标记发生倍增并转移到别的染色体臂上从而产生了沉默区域。端粒处的组成型异染色质以染色质“帽”的形式保证了基因组的稳定性。最后，异染色质形成还是一种针对入侵 DNA 的防卫机制。总之，这些发现强调了一个重要的观点，即，异染色质对基因组的维持有着重要的功能，其重要性是可以和常染色质相提并论的。

总体而言，常染色质和异染色质之间的广泛的功能差异可以归结于染色质的三个已知特征。第一，DNA 序列本身。例如，是否含有启动子周围的 AT 富集的“刚性”DNA、重复序列和（或）阻遏因子结合序列。第二，转录生成的 RNA 的质量，这决定了它是被加工成成熟的 mRNA 用于翻译，或者是被降解，还是被用于引导 RNAi 机器实现其靶序列的异染色质化。第三，核内的空间组织，这起着重要的隔离作用，用来维持该处的染色质构象。

## 7. 组蛋白修饰与组蛋白密码

我们已经探讨了组蛋白修饰是如何改变染色质模板的：或是通过顺式作用，改变核小体间的接触和间距；或是通过与染色质模板相结合的组蛋白和非组蛋白产生的反式作用。但组蛋白修饰的功能及生物学效应是什么呢？对总组蛋白的研究表明组蛋白尾部修饰与染色质的结构模式具相关性，这暗示表观遗传标记可能提供了“开”（即：活化）或“关”（沉默）的信号（见附录）。这是从长期的、大多为关联性的研究中产生的，这些研究表明，一些组蛋白修饰，特别是组蛋白乙酰化与通常允许转录的活性染色质区域相关。相反，其他标记，如某些磷酸化修饰的组蛋白残基多与通常不利于转录活性的致密态染色质相关。附录 B 列出了现如今已知的组蛋白修饰位点。在此我们强调这些修饰及位点并不一定存在于所有生物中。

首先，组蛋白修饰是如何被确立或去除的呢？染色质领域的大量研究表明，组蛋白尾部修饰的确立（“写入”）或去除（“擦去”）是通过染色质相关酶系统的催化作用完成的。然而，这些酶在许多年中一直未被鉴定出来。直至过去十年，科学家们从多种来源鉴定出大量的染色质修饰酶，它们中的大部分已被编纂在附录 B 里。这是通过大量的生物化学及遗传学研究得到的。这些酶常存在于大的多亚基复合物中，催化组蛋白与非组蛋白底物上共价修饰的掺入或去除。此外，某些酶的催化活性具有显著的靶向残基特异性及细胞环境（即依赖于胞外及胞内信号）特异性。为了更清楚地阐明这个问题，我们将举例简要讨论催化组蛋白修饰的四种主要的酶系统及其对应的逆向修饰酶系统（图 3-10；Vaquero et al. 2003；Holbert and Marmorstein 2005）。这些拮抗活性共同控制

了我们讨论的每一种修饰的稳态平衡状态。

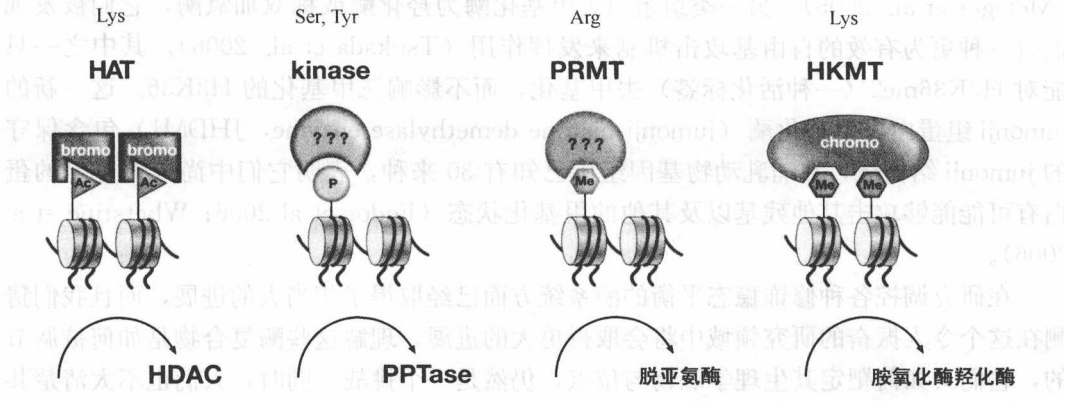



图 3-10　组蛋白修饰酶

组蛋白共价修饰由组蛋白修饰酶进行“写手”催化，去除其拮抗活性。根据酶促反应类型（如乙酰化或磷酸化）的不同，它们被列入不同的家族。与组蛋白尾部某种修饰有特异性亲和力的蛋白质结构域被称作“读者”。HAT：组蛋白乙酰化酶；PRMT：蛋白精氨酸甲基转移酶；HKMT：组蛋白赖氨酸甲基转移酶；HDAC：组蛋白去乙酰化酶；PPTase：蛋白磷酸酶；Ac：乙酰化；P：磷酸化；Me：甲基化。

组蛋白乙酰化酶（Histone acetylase，HAT）催化组蛋白底物中特定赖氨酸残基的乙酰化（Roth et al. 2001），这种乙酰化作用能被组蛋白去乙酰化酶（histone deacetylase，HDAC）所逆转（Grozinger and Schreiber 2002）。组蛋白激酶家族的成员能够使特定的丝氨酸或苏氨酸残基磷酸化，而磷酸酶则可以去除这些磷酸基团标签。值得一提的是，众所周知的有丝分裂激酶，如 cyclin 依赖的激酶或 aurora 激酶，能够催化核心组蛋白（H3）及连接组蛋白（H1）的磷酸化。细胞脱离有丝分裂状态时，逆转磷酸化作用的磷酸酶（opposing phosphatase，PPTase）都还不明了。

已经报道的甲基化酶主要有两类：组蛋白精氨酸甲基转移酶（protein arginine methyltransferase，PRMT）的底物是精氨酸残基（Lee et al. 2005），组蛋白赖氨酸甲基转移酶（histone lysine methyltransferase，HKMT）作用于赖氨酸残基（Lachner et al. 2003）。精氨酸瓜氨酸化酶能将甲基化精氨酸（或精氨酸）转变为瓜氨酸残基，从而间接地逆转精氨酸甲基化（Bannister and Kouzarides 2005）。甲基化的赖氨酸残基看起来具有较高的化学稳定性。赖氨酸的甲基化状态有单甲基化、双甲基化和三甲基化三种形式。H3 和 H4N 端的一些三甲基化残基似乎在细胞分裂过程中具有稳定遗传的能力（Lachner et al. 2004），类似的有果蝇成虫盘中的 H4K20me1 标记（Reinberg et al. 2004）。最近发现，一种被称作赖氨酸特异性“去甲基化酶”（lysine specific demethylase，LSD1）的胺氧化酶能去除 H3K4 的甲基化修饰（Shi et al. 2004）。这种酶通过 FAD 依赖的氧化来破坏氨-甲基键，最终产生非修饰的赖氨酸和甲醛。据报道，LSD1 对激活性 H3K4 甲基化标记具有选择性，它只能作用于单甲基化和双甲基化的 H3K4，而不作用于三甲基化 H3K4。这种“去甲基化酶”是大的抑制性蛋白复合物的组分，这一蛋白复合物还包括 HDAC 和其他一些酶。另有证据显示：LSD1 与雄激素受体在靶

位点形成一个复合物，通过对抑制性组蛋白标记 H3K9me2 去甲基化来促进转录激活（Metzger et al. 2005）。另一类组蛋白去甲基化酶为羟化酶或称双加氧酶，它们被发现通过一种更为有效的自由基攻击机制来发挥作用（Tsukada et al. 2006）。其中之一只能对 H3K36me2（一种活化标签）去甲基化，而不影响三甲基化的 H3K36。这一新的 jumonji 组蛋白去甲基化酶（jumonji histone demethylase enzyme，JHDM1）包含保守的 jumonji 结构域，在哺乳动物基因组中已知有 30 来种，表明它们中尚未被鉴定的蛋白有可能能够攻击其他残基以及其他的甲基化状态（Fodor et al 2006；Whetstine et al 2006）。

在研究调控各种修饰稳态平衡的酶系统方面已经取得了相当大的进展，而且我们猜测在这个令人振奋的研究领域中将会取得更大的进展。理解这些酶复合物是如何被调节的，它们又如何靶定其生理学底物与位点，仍然是一个挑战。同时，人们也不太清楚共价机制是如何影响表观遗传现象的。

组蛋白修饰并不是独立发生的，而是组合成所谓的修饰盒（即在特定的组蛋白尾部的邻近残基发生多个共价修饰，例如 H3K9me 和 H3S10ph、H4S1ph、H4R3me 和 H4K4ac）和反式组蛋白通路（不同组蛋白尾巴或核小体间的共价修饰组合，图 3-11）。有趣的是，大部分已知组蛋白修饰依其修饰的组蛋白 N 端氨基酸残基位点的不同，而与激活或抑制功能具相关性。协同及拮抗通路均已被描述（Zhang and Reinberg 2001；Berger 2002；Fischle et al. 2003b），它们能够逐渐诱导活化标记的协同作用，并同时抵消抑制性修饰。然而，由于大部分的研究以总组蛋白为材料，因此，还不知道在给定核小体中存在多少种不同组蛋白 N 端上的不同修饰的组合。除 N 端外，组蛋白球状折叠结构域的修饰也已被证明能影响染色质的结构和组装（Cosgrove et al. 2004），进而影响基因表达和 DNA 损伤修复（van Attikum and Gasser 2005；Vidanes et al. 2005）。此外值得注意的是，一些组蛋白修饰酶能够作用于非组蛋白底物（Sterner and Berger 2000；Chuikov et al. 2004），图 3-11 列举了两例已经明确的组蛋白修饰层级，似乎可以作用于对活跃染色质的转录检索，亦或反之，指导异染色质区域的模式化。

这些研究提出了一个问题：是否存在“组蛋白密码”或者是“表观遗传密码”？尽管这个理论概念极具诱惑力，且其某些预言已被证实，然而，“密码”是否真正存在仍是一个未解之谜。相比较而言，遗传密码具可预测性和通用性，被证明极为有用。绝大多数情况下，它使用 DNA 中的四个基本“字母”（即核苷）形成了一种通常不变的，近乎通用的语言。相反，目前的证据表明，组蛋白修饰模式随物种而变化，尤其是在“低等”和“高等”的真核生物，例如酵母和人类之间。因此，即使存在组蛋白密码，它也不太可能具有通用性。当考虑到组蛋白修饰依据空间和时间变化的动态性，这就变得更加复杂。此外，染色质模板还存在大量的重塑因子（Vignali et al. 2000；Narlikar et al. 2002；Langst and Becker 2004；Smith and Peterson 2005）。但是，当染色质免疫沉淀实验（chromatin immunoprecipitation assays，ChIP）被运用到全基因组水平时（ChIP on chip），一些基因组（如 *S. pombe*，*A. thaliana*，哺乳动物细胞）中非随机的和某种程度上可预测的修饰模式开始得以破译，例如 H3K4me3 与活化启动子区（Strahl et al. 1999；Santos-Rosa et al. 2002；Bernstein et al. 2005）及 H3K9（Hall et

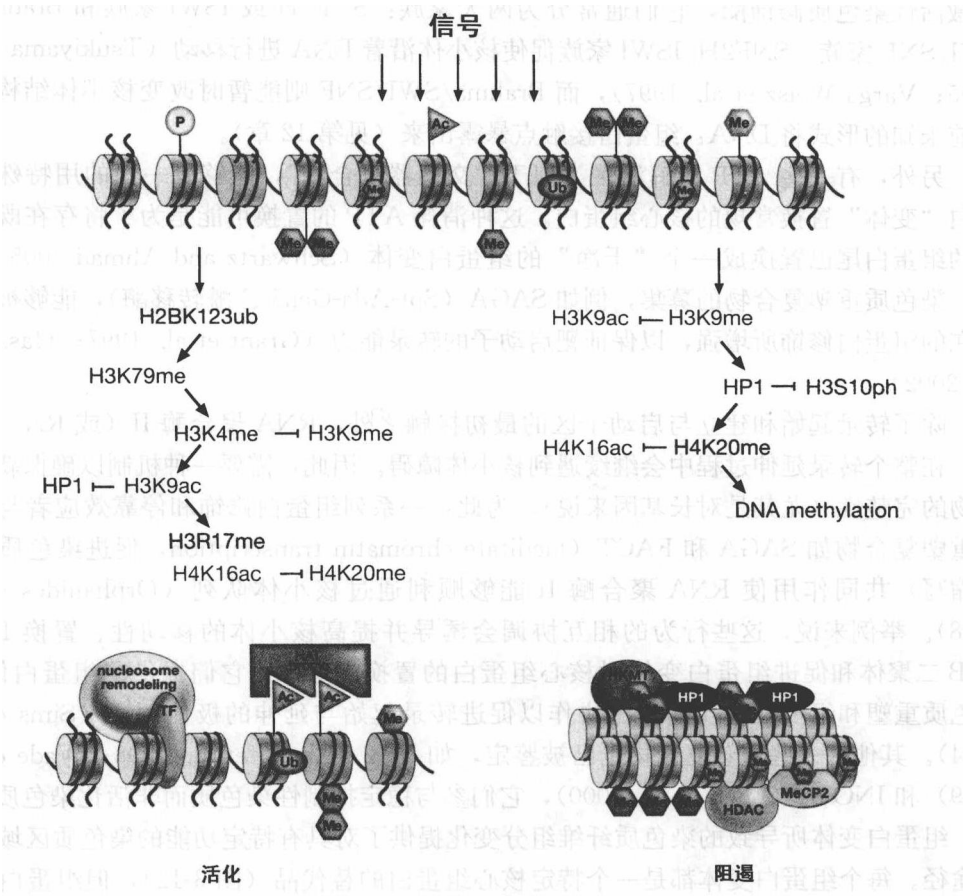



图 3-11　染色质的协同修饰

非活化染色质模板到活性常染色质的转变（左）或抑制性异染色质的确立（右）包含了一系列协同的染色质修饰。在转录激活的例子中，伴随了核小体重塑复合物的作用和组蛋白变体对核心组蛋白的置换。

al. 2002；Lippman et al 2004；Martens et al. 2005）和 H3K27（Litt et al. 2001；Ringrose et al. 2004）甲基化与沉默异染色质的强相关性。也许，组蛋白密码的局限性在于一种修饰不能恒定地转换成一种生物信号。然而，组合或累积的修饰似乎确实能够定义和实现某种生物学功能（Henikoff 2005）。

## 8. 染色质重塑复合物与组蛋白变体

诱导染色质模板发生转换的另一个主要的机制是募集染色质“重塑”复合物，以一种非共价的形式，利用能量（ATP 水解酶）来改变染色质和核小体的构象。核小体，尤其被抑制性染色质结合因子结合时，常对转录机器构成天然的障碍。因此，只有一些序列特异性转录因子和调控因子（而非转录基本机器）才能够接近它们的结合位点。这一问题部分地由可移动或改变核小体结构的蛋白复合物所解决。染色质重塑活性常伴随

着激活性染色质修饰酶，它们通常分为两大家族：SNF2H 或 ISWI 家族和 Brahma 或 SWI/SNF 家族。SNF2H/ISWI 家族促使核小体沿着 DNA 进行移动（Tsukiyama et al. 1995；Varga-Weisz et al. 1997），而 Brahma/SWI/SNF 则能暂时改变核小体结构，以目前未知的形式将 DNA：组蛋白接触点暴露出来（见第 12 章）。

另外，有一些 ATP 水解活性类似于“交换器复合物”，能够专一性的用特殊的组蛋白“变体”置换常规的核心组蛋白。这种消耗 ATP 的置换可能是为了将存在既有修饰的组蛋白尾巴置换成一个“干净”的组蛋白变体（Schwartz and Ahmad 2005）。此外，染色质重塑复合物的募集，例如 SAGA（Spt-Ada-Gcn5-乙酰转移酶），能够被已经存在的组蛋白修饰所增强，以保证靶启动子的转录能力（Grant et al. 1997；Hassan et al. 2002）。

除了转录起始和建立与启动子区的最初接触之外，RNA 聚合酶 II（或 RNA 聚合酶）在整个转录延伸过程中会继续遇到核小体障碍。因此，需要一种机制以确保新生转录物的完整性（尤其是对长基因来说）。为此，一系列组蛋白修饰和停靠效应者与染色质重塑复合物如 SAGA 和 FACT（facilitate chromatin transcription，促进染色质转录的缩写）共同作用使 RNA 聚合酶 II 能够顺利通过核小体队列（Orphanides et al. 1998）。举例来说，这些行为的相互协调会诱导并提高核小体的移动性、置换 H2A/H2B 二聚体和促进组蛋白变体对核心组蛋白的置换。这样，它们提供了组蛋白修饰、染色质重塑和组蛋白置换间紧密协作以促进转录起始与延伸的极佳例证（Sims et al. 2004）。其他的一些重塑复合物也已被鉴定，如 Mi-2（Zhang et al. 1998；Wade et al. 1999）和 INO-80（Shen et al. 2000），它们参与稳定抑制性染色质而非活化染色质。

组蛋白变体所导致的染色质纤维组分变化提供了对具有特定功能的染色质区域的检索途径。每个组蛋白变体都是一个特定核心组蛋白的替代品（图 3-12），但组蛋白变体在总组蛋白中仅占极少一部分，因此，相对于一般的组蛋白而言，对组蛋白变体的研究更为困难。越来越多的文献（综述见 Henikoff and Ahmad 2005；Sarma and Reinberg 2005）表明，组蛋白变体有着独特的修饰模式，这很可能是因为它们的少数氨基酸变化使之有别于其他家族成员。此外，一些组蛋白变体具有不同的 N 端和 C 端结构域，使之具有独特的染色质调节活性和与结合因子的不同亲和力。举例来说，转录活跃基因的组蛋白 H3 一般会被 H3.3 变体以一种转录偶联的机制置换掉，而不依赖于 DNA 的复制（Ahmad and Henikoff 2002）。核心组蛋白 H2A 被 H2A.Z 的替换与转录活性相关，并能指示出无核小体启动子的 5′端。然而，H2A.Z 也存在于抑制性染色质。CENPA 是一种着丝粒特异性的组蛋白变体，它对于着丝粒的功能也即染色体分离至关重要。H2A.X 与其他组蛋白标记共同感知 DNA 损伤并指引 DNA 修复复合物富集到 DNA 损伤处。MacroH2A 与哺乳动物中 X 染色体失活（inactive X chromosome，Xi）相关的一种特异性组蛋白变体（关于组蛋白变体的更多详细信息见第 13 章）。

重要的是，与被广为认可的组蛋白合成及包装于 S 期的教科书观点相反，许多组蛋白变体的合成及替代不依赖于 DNA 复制。因此，组蛋白变体对核心组蛋白的置换不受细胞周期所处阶段（即 S 期）的限制，而是能对正在发生的事件（例如，转录活性或细胞分裂过程中的着丝粒牵引）或“胁迫信号”（例如，DNA 损伤或营养饥饿）进行及时

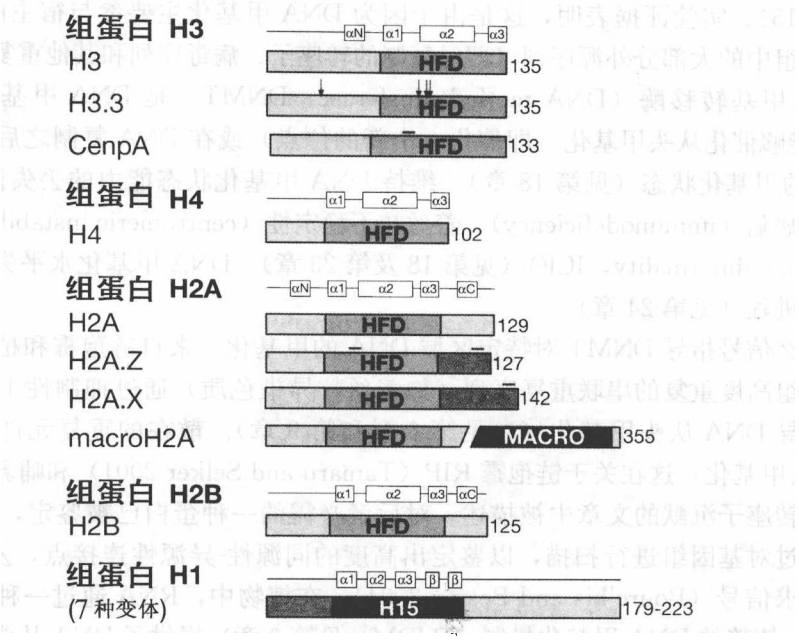



图 3-12　组蛋白变体

核心组蛋白（H3，H4，H2A，H2B）、连接组蛋白 H1 以及组蛋白 H3 和 H2A 变体的结构域。HFD：histone fold domain，组蛋白折叠结构域。

的响应。精心设计的生化研究证明染色质重塑或交换复合物特异性地参与了不同组蛋白变体例如 H3. 3、H2A. Z 或 H2A. X 的置换（Cairns 2005；Henikoff and Ahmad 2005；Sarma and Reinberg 2005）。例如，H3. 3 变体对 H3 的置换是通过 HIRA（组蛋白调节子 A）交换复合物的作用进行的（Tagami et al. 2004）；H2A. Z 对 H2A 的置换是通过 SWR1（Swi2/Snf2-related ATPase 1）交换复合物的作用进行的（Mizuguchi et al. 2004）。这些置换作用共同使核小体变体构建出特别活跃的染色质。对于其他的一些组蛋白变体来说，靶向及交换是通过 ATP 依赖的组蛋白交换复合物进行的，还是通过组蛋白伴侣蛋白进行的还需要进行进一步的探索。人们推测，交换复合物能够用非修饰的组蛋白替换已被修饰的组蛋白，以去除组蛋白 N 端更加稳定的表观遗传标记。

## 9. DNA 甲基化

DNA 甲基化是已知最早被发现的与基因抑制相关的表观遗传机制（Razin and Riggs 1980）。它在酵母以外的所有真核生物中都有不同程度的存在。这种修饰是在 DNA 模板的胞嘧啶上加上甲基基团。在哺乳动物中，它发生于 CpG 二核苷酸，而在链孢霉和植物中存在其他对称的、非对称的和非 CpG 的甲基化模式。甲基化 DNA 在基因组上的分布表明它们集中出现在非编码区（如着丝粒异染色质）和散在的重复元件（转座子）区域，而不是活化基因的“CpG 岛”区（Bird 1986）。事实上，DNA 甲基化水平提高的程度是与高等真核细胞基因组中非编码区和重复 DNA 序列的相对增加相关

的（图 3-15）。实验证据表明，这是由于因为 DNA 甲基化主要参与宿主的防御机制：沉默基因组中的大部分外源序列（即经复制的转座子、病毒序列和其他重复序列）。

DNA 甲基转移酶（DNA methyltransferase，DNMT）是 DNA 甲基化的“效应器”，它能够催化从头甲基化（即催化一个新的位点）或在 DNA 复制之后维持半甲基化 DNA 的甲基化状态（见第 18 章）。维持 DNA 甲基化状态能力的丢失能导致疾病，例如免疫缺陷（immunodeficiency）、着丝粒不稳定性（centromeric instability）和面部畸形（facial abnormality，ICF）（见第 18 及第 23 章）。DNA 甲基化水平失调也参与肿瘤发生的进程（见第 24 章）。

是什么信号指导 DNMT 对特定区域 DNA 的甲基化？来自链孢霉和植物的证据显示，基因组高度重复的串联重复序列（如着丝粒异染色质）通过抑制性 H3K9 甲基化标签来指导 DNA 从头甲基化（参见第 6 章和第 9 章）。散在的重复元件也能够指导 DNA 从头甲基化，这在关于链孢霉 RIP（Tamaru and Selker 2001）和哺乳动物雄性生殖系中逆转座子沉默的文章中被描述。对后者必需的一种蛋白已被鉴定，即 Dnmt3L，它可能通过对基因组进行扫描，以鉴定出高度的同源性-异源性连接点，发出 DNA 甲基化的需求信号（Bourc'his and Bestor 2004）。在植物中，RNA 通过一种独特的，被称为 RNA 依赖的 DNA 甲基化机制（RdDM；见第 9 章）提供了 DNA 从头甲基化的信号。有证据显示，SWI/SNF 家族的染色质重塑因子在某些方面对于维持 DNA 甲基化的总体模式是必须的，如植物中 DDM1 蛋白（Jeddeloh et al. 1999）和其哺乳动物中的同源蛋白 Lsh1（Yan et al. 2003）。最后，HKMT-PcG 蛋白 Ezh2 也可能参与指导了哺乳动物中的 DNA 甲基化（Vire et al. 2005）。

当 DNA 甲基化被建立后，它通过何种机制作用于染色质沉默还不完全清楚，尽管有证据表明是通过反式调节作用进行的。甲基化胞嘧啶的结合因子被称为甲基化 CpG 结合结构域蛋白（methyl-CpG-binding domain proteins，MBD），被认为在 DNA 甲基化现象中等同于组蛋白修饰结合因子（图 3-13）。例如，甲基胞嘧啶结合蛋白（methyl-cytosine-binding protein，MeCP2）能够结合甲基化 CpG 并募集 HDAC 来介导抑制性组蛋白标签（见第 18 章）。另外，DNA 甲基化能干扰转录调节因子的识别位点（如 CTCF，见第 19 章基因组印记）。

基因印记区域甲基化 DNA 的存在导致植物中和胎盘类哺乳动物中母本或父本等位基因的沉默，包括 X 染色体失活，这表明在进化过程中它们独有地利用了这一表观遗传机制来稳定沉默。有趣的是，有袋类动物中，基因印记区域没有 DNA 甲基化，这表明甲基化 DNA 对胎盘类哺乳动物中基因印记区域的沉默是相对较近的进化事件（见第 17 章和第 19 章）。与之相反的是，双翅类昆虫例如果蝇中，DNA 甲基化作为一种功能性表观遗传机制已基本被丢失（Lyko 2001）。

当哺乳动物基因组中通常甲基化的高度重复区域处于非甲基化状态时，会逐步变得具有诱变性，乃至影响基因组稳定性（Chen et al. 1998）。继而发生的染色质异常是许多疾病和肿瘤发生的主要原因（见第 15 节）。这凸显了 DNA 甲基化在维持基因组完整性中的重要作用。相反的是，单个胞嘧啶碱基的甲基化具有高度的自发性突变倾向。随时间推移，通过脱氨基反应导致从 C 到 T 的转换（图 3-13）。但这种特性被认为有利于

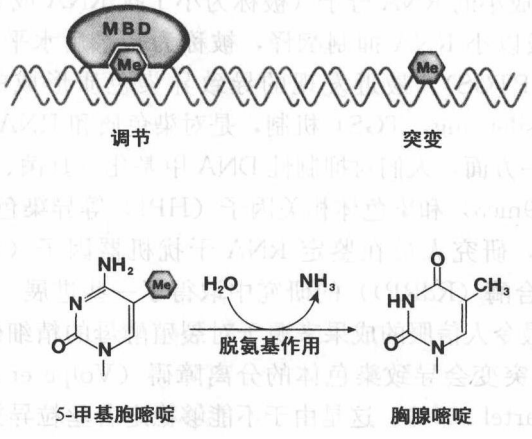



图 3-13　DNA 甲基化和脱氨基化

甲基化（Me）的胞嘧啶核苷酸能够被甲基化 DNA 结合蛋白（methyl DNA-binding proteins，MBD）所结合。未被结合的 5-甲基胞嘧啶易通过脱氨基化（下图）自发产生突变，使 DNA 序列中的 5-甲基 CpG 转变为 TpA。

宿主基因组的保护，因为它能够永久地抑制寄生性 DNA 序列，例如转座子。在不同的背景下，相同的化学反应可以被“活化诱导的脱氨基酶”（AID）所催化。该酶在 B 细胞和 T 细胞中的表达导致免疫球蛋白（Ig）基因座的“体细胞高突变”。这对于丰富抗原受体库，从而增强哺乳动物免疫力是一个重要的机制（第 21 章；Petersen-Mahrt 2005）。AID 在哺乳动物发育早期的表达导致一种设想：尽管存在点突变率增加的风险，它的存在仍有可能给 DNA 去甲基化提供了一条可供选择的途径。

DNA 甲基化和组蛋白甲基化是基因组表观遗传调控的代表性机制。下一节描述的非编码 RNA 是染色质沉默的重要的最初诱因。RNA 分子的糖基或核苷酸骨架也已知能被高度甲基化。此外，小的非编码 RNA 3′末端的甲基化已被证实能够稳定这些分子（Yu et al. 2005）。因此，就像 DNMT 和 HKMT 一样，RNA 甲基转移酶也可能是表观遗传信息的“作者”，尽管现在还没有直接的证据。然而，RNA 甲基化似乎能被一些类 Toll 受体（识别病原子共同模式的跨膜受体）所感知，并指导先天性免疫（Ishii and Akira 2005），这一例证支持上述假说。这就产生了一种有趣的可能性：RNA 甲基化可能是有待证明的第三种形式的，基于甲基化的表观遗传调节方式。

## 10. RNA 干扰与 RNA 指导的基因沉默

对着丝粒及端粒处组成型异染色质在维持基因组稳定性中的重要作用的了解，导致了人们对重复性非编码“垃圾”DNA 的看法的发生了根本性转变。这些重复序列可能并非毫无用处，只是其功能刚刚开始将被阐明？甚至这种 DNA 序列并非完全处于“沉默”状态？

这种可能性来源于将 RNA 干扰与沉默染色质（异染色质）的形成相联系的一系列基本发现。RNA 干扰是宿主的一种防御机制，它能够将双链 RNA（double stranded

RNA，dsRNA）切割成小的 RNA 分子（被称为小干扰 RNA 或 siRNA）。这个过程最终导致 RNA 的降解或以小 RNA 抑制翻译，被称为转录后水平基因沉默（post-transcriptional silencing，PTGS）。最近发现的导致异染色质形成的转录水平基因沉默（transcriptional gene silencing，TGS）机制，是对染色质和 RNA 干扰机制两者的独立研究交汇后发现的。一方面，人们对抑制性 DNA 甲基化（真菌、植物和哺乳动物中）、染色质修饰（如 H3K9me3）和染色体相关因子（HP1）等异染色质区域的一些特性了解了很多。另一方面，研究人员在鉴定 RNA 干扰机器因子（如 Dicer、Argonaute、RNA 依赖的 RNA 聚合酶（RdRP））的研究中取得了一些进展。将两个看似毫不相关的领域联系到一起的最令人信服的成果来源于对裂殖酵母的精细研究，在这种酵母中，RNA 干扰机器组分的突变会导致染色体的分离障碍（Volpe et al. 2002；Hall et al. 2002；Reinhart and Bartel 2002）。这是由于不能够稳定着丝粒异染色质而导致的结果，凸显了 RNA 干扰介导的机制在沉默异染色质区域的形成过程中可能存在普遍性。着丝粒异染色质对于染色体的分离过程是必须的，这也突出了异染色质在转录沉默基因以外，对维持基因组完整性及生存方面的重要性。也有证据表明 siRNA 在定义其他功能异染色质，例如端粒的特定区域中是必须的。

dsRNA 衍生物是促使 RITS 复合物识别着丝粒并将其沉默的关键底物，裂殖酵母着丝粒周边重复 DNA 序列两条链的转录和加工后的 siRNA 衍生物的发现为此提供了强有力的证据（图 3-14）（Verdel et al. 2004）。此外，clr4 突变体（裂殖酵母中哺乳动物组蛋白甲基转移酶 Suv39h 的同源物）不能将 dsRNA 加工成 siRNA，这进一步强调了 RNA 干扰机器与异染色质组装之间存在着相互影响的事实（Motamedi et al. 2004）。至今人们还不太清楚由 RNA 干扰机器（即 Dicer、Argonaute 和 RdRP）产生的 siRNA 是如何正确起始异染色质组装或者如何“指导”siRNA 进入一个合适的基因组区域的。由此，人们提出一个模型：RNA 干扰机器复合物 RITS 与着丝粒重复序列之间的相互作用导致异染色质的形成进入一个自增强循环路径，其中有 Clr4、HDAC、Swi6（哺乳动物 HP1 的同源物）和 cohesin 的参与，并可能通过 Ago 介导的 RNA：RNA 杂交与新生转录物配对的方式来实现（图 3-14）（见第 6 章和第 8 章）。

在一种有纤毛的原生动物嗜热四膜虫中，人们发现一种类似的 RNA 介导的靶定机制，它能指导体细胞核内发生的 DNA 剔除这一独特事件。在这个事例中，在性别途径的适当阶段，嗜热四膜虫“沉默”的、种系（微核）基因组中的内部剔除元件（internal eliminated segment，IES）发生 DNA 双链的转录（Chalker and Yao 2001；Mochizuki et al. 2002）。与 RNA 干扰依赖的 TGS 模型相仿，“扫描 RNA”（scan RNA，scnRNA）模型被提出用于解释亲本滋养核中的 DNA 序列是如何在新的滋养核中表观调控基因组的改变，小 RNA 参与了这一过程（详见第 7 章）。这些令人兴奋的结果初次提供了 RNA 类似进程能够直接改变体细胞基因组的证据。由此产生了一种有趣的可能性：V-DJ 处产生的基因间 RNA（Bolland et al. 2004）有可能在 B 细胞 IgH 基因座和 T 细胞 TCR 基因座的 V-DJ 重排过程中指导了 DNA 序列的剔除。

在植物中，许多 RNA 干扰机器组分具有众多同源蛋白，导致一系列 RNA 沉默途径，尽管这些因子中存在冗余现象，但是它们的产生可能对特定的 DNA 序列具有更强

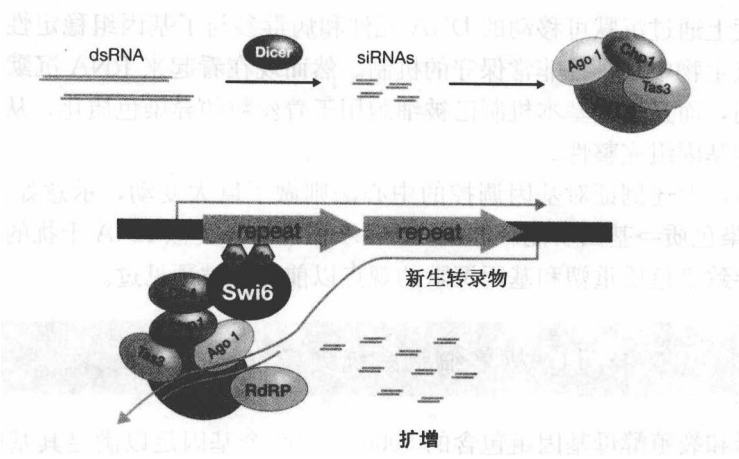



图 3-14　RNA 指导的异染色质形成

由两条 DNA 链转录或由反向重复转录物折叠而产生的互补 dsRNA 转录物在 Dicer 的作用下产生 siRNA。siRNA 通过与 Argonaute（Ago）蛋白的结合整合到 RITS 复合物中，激活该复合物以识别互补 DNA 或新生 RNA。这个复合物招募 Clr4，产生组蛋白 H3K9me2。修饰特异性的结合因子 Swi6 与含有这些修饰的组蛋白结合，进而促进抑制性染色质区域的扩展。RdRP 以现存的 siRNA 为引物扩增 siRNA，并增强 RITS 复合物对特定 DNA 区域的识别能力。

的专一性。植物中 RNA 干扰介导的 TGS 的研究揭示了一类新的 RNA 聚合酶——RNA 聚合酶 IV（或 RNA Pol IV），它只转录异染色质区域 DNA（Herr et al. 2005；Pontier et al. 2005）。植物中另一特有的现象是 RNA 干扰途径直接影响 DNA 的甲基化（Chan et al. 2004；详见第 9 章）。

在果蝇和哺乳动物中也发现了类似 RNA 干扰的染色质效应。例如用 RNAse A 作用于可渗透的哺乳动物细胞迅速去除异染色质的 H3K9me3 标签，表明某种 RNA 组分可能是着丝粒周围异染色质的一个结构成分（Maison et al. 2002）。脊椎动物中 siRNA 加工因子的去除削弱了 H3K9 的甲基化水平和 HP1 与着丝粒周围异染色质的结合（Fukagawa et al. 2004）。有趣的是，dicer$^{-/-}$ 的突变体胚胎干细胞（embryonic stem，ES）能继续增殖，但丧失了分化能力（Kanellopoulou et al. 2005），这表明 RNA 干扰机器、非编码 RNA 和哺乳动物发育之间存在一种尚未被了解的联系。在果蝇中，mini-white 基因串联矩阵的沉默，或位置效应花斑（position effect variegation，PEV）的主体，也依赖于 RNA 干扰机器（Pal-Bhadra et al. 2004）。

综上所述，这些研究说明了非编码 RNA 在引发表观遗传转换和可继承地维持染色质模版的特定染色质状态中具有重要的，甚至可能是最根本性的角色。事实上，这些非编码 RNA 为不同生物体中众多不同的重复序列如何通过 RNA 靶定机制实现异染色质化做出了解答。在植物、果蝇、哺乳动物和其他物种中，为鉴定更多的 RNA 干扰靶点，通过对小 RNA 测序，揭示了它们大部分由内源转座子和其他重复序列转录而来

(Bernstein and Allis 2005; Almeida and Allshire 2005)。总之，这些结果表明 RNA 干扰在一定程度上通过沉默可移动的 DNA 元件和病毒参与了基因组稳定性的维持，并且在大多数真核生物中是一种非常保守的机制。然而现在看起来 RNA 沉默不仅仅抑制了外源入侵序列，而且这种基本机制已被细胞用于着丝粒的异染色质化，从而确保正确的染色体分离和基因组完整性。

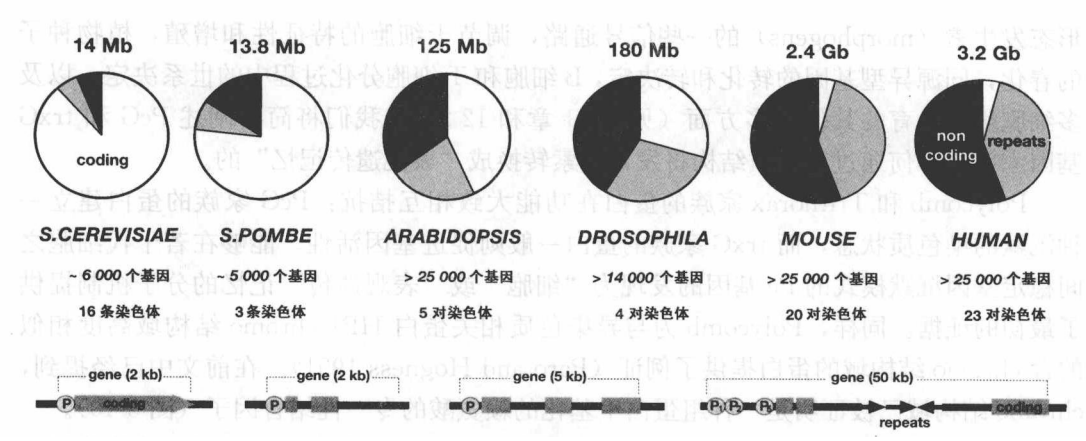
总体而言，上述例证对基因调控的中心法则做了巨大变动，示意如下：DNA⇒非编码 RNA⇒染色质⇒基因功能。非编码 RNA 积极参与类似 RNA 干扰的机制，并作用于特定区域导致染色质重塑和基因沉默的观点以前从未被预见过。

## 11. 从单细胞系统到多细胞系统

酿酒酵母和裂殖酵母基因组包含的 5000～6000 个基因足以满足其基础代谢和细胞分裂过程的需要。然而，由于这些单细胞生物基本上是无性繁殖的，仿佛永远不朽的生命体，因此对细胞分化没有要求。相比之下，哺乳动物细胞需要编码约 25 000 个基因以满足约 200 种不同细胞类型的需要。因此，理解如何从一套相同的遗传模版中产生并协调多细胞的复杂性是表观遗传学研究的一个关键问题。

通过对酵母、果蝇、植物和哺乳动物的基因组进行比较，发现基因组的大小是随着各物种的复杂程度而显著扩增的。酵母和哺乳动物基因组的大小相差超过 300 倍，但是基因数量却仅仅增加了 4～5 倍（图 3-15）。但是，编码区与非编码区和重复序列的比值表明了基因组的复杂性：与多细胞生物高度异染色质化的基因组相比，单细胞真菌类的基因组绝大多数处于“开放”状态，基本上没有什么非编码区。尤其是哺乳动物积聚有大量的重复元件和非编码区，它们占据了大部分 DNA 序列（52%非编码区和 44%重复 DNA）。因此，哺乳动物只有 4%的基因组用于编码功能蛋白（包括内含子序列）。多细胞生物中重复序列及非编码区的大量增加很有可能是由于整合了入侵的外来元件，例如 DNA 转座子、逆转座子和其他重复元件。尽管它们给基因表达程序的协调性带来负担，但它们也使基因组得以进化并保持其可塑性，以及一定程度的随机的基因调节。重复元件的扩增甚至渗透到了哺乳动物基因组的转录单元，使得转录单元经常变得更大（30～200 kb），且其非翻译的内含子区常常有多重启动子和 DNA 重复序列的存在。相反，虽然植物也拥有相似的大基因组，但是它们一般只有较小的转录单元和内含子，这是因为它们进化出一套防御机制以保证其转录单元内不会出现转座子的插入。

虽然不同物种中一些作用机制具有高度的功能保守性，但不同物种所用的表观遗传通路类型有着重要的物种差异性。在一定程度上，这种差异与基因组的大小有关。高等真核细胞基因组的高度扩增（包括非编码区和重复 DNA 序列）需要更为广泛的表观遗传沉默机制。相应的，在哺乳动物与植物中存在大量的阻遏性组蛋白赖氨酸甲基化、DNA 甲基化和 RNA 干扰沉默机制。伴随多细胞生物的另一挑战是如何协调和维持细胞类型（细胞特征性）的多样性。这是一个精确的平衡，包括聚梳（PcG）和三胸（trxG）家族的蛋白复合物对基因组的调节，特别是聚梳蛋白与多细胞的出现相关联（参见第 12 章）。




**简单与复杂的基因组成**

图 3-15　不同生物基因组组成分布图

每个饼状图的上方列出了表观遗传学研究的主要模式生物的基因组大小。在复杂的多细胞生物中，基因组大小的增加是与非编码区（即内含子、基因间和散在的重复序列）和重复 DNA（如卫星 DNA、长散布元件和短散布元件）序列的大量扩增相关联的。这种扩增伴随着调控基因组的表观遗传机制（特别是抑制性的）数量的增加。同时，基因组的扩增还伴随着转录单元大小及复杂度的增加。植物除外，因为它们演化出了一套阻止在转录单元内部进行插入或倍增的机制。

多细胞生物的细胞从功能上可以被分为两大类：生殖细胞（全能的，为遗传信息传递到下一代所必需）和体细胞（分化的，承担各种生命功能的生物“工厂”）。生殖细胞如何保持其表观基因组的全能性，拭去、建立和保持细胞命运（“细胞记忆”）的机制又是什么，这些都是非常重要的问题。由于一个生殖细胞能产生另一个生殖细胞，本来它具有无限的生命潜力，就像单细胞的“不朽”生命体那样。然而，为了完成这个角色，生殖细胞在绝大多数时间处于“静息”状态，不响应外界刺激，以保护其表观基因组的完整性。事实上，哺乳动物卵母细胞能够保持其静息状态超过 40 年。类似的，大多数成体干细胞（多能性的）也多处于休眠状态，只有在被有丝分裂刺激信号激活后，才进入次数有限的细胞分裂过程进行增殖（或称自我更新）。因此，表观基因组的组成经受着许多内在（如：转录、DNA 复制、染色质分离）和外在（如：细胞因子、激素、DNA 损伤或一般说来的“胁迫”作用）信号的挑战，尤其是当体细胞分化迫使其离开具有保护作用的生殖细胞和干细胞环境时。

## 12. 聚梳和三胸

能够将信号传递到染色质模版并参与维持细胞特征性（即提供细胞记忆）的主要效应器中有些是聚梳（Polycomb，PcG）和三胸（Trithorax，trxG）基因家族的成员（Ringrose and Paro 2004）。这些基因因其在果蝇 *Hox* 基因簇发育调节及同源异型基因调节中的作用而被发现。从此，PcG 和 trxG 被证明是多细胞真核生物中细胞增殖和细胞特征性的关键调节因子。此外，这类基因还参与了应答促有丝分裂原（mitogens）和

形态发生素（morphogens）的一些信号通路，调节干细胞的特征性和增殖，植物种子的春化，同源异型基因的转化和转决定，B 细胞和 T 细胞分化过程中的世系决定，以及多细胞生物发育的其他许多方面（见第 11 章和 12 章）。我们将简要阐述 PcG 和 trxG 基因家族是如何通过染色质结构将发育线索转换成“表观遗传记忆”的。

Polycomb 和 Trithorax 家族的蛋白在功能大致相互拮抗：PcG 家族的蛋白建立一种沉默的染色质状态，而 trxG 家族的蛋白一般则促进基因活性。能够在若干代细胞之间稳定基因沉默模式的 Pc 基因的发现为“细胞”或“表观遗传”记忆的分子机制提供了最初的证据。同样，Polycomb 为与异染色质相关蛋白 HP1 chromo 结构域高度相似的含 chromo 结构域的蛋白提供了例证（Paro and Hogness 1991）。在前文中已经提到，chromo 结构域已被证明是一种组蛋白甲基化的赖氨酸的专一性结合因子（图 3-10）。

在果蝇中已经鉴定出约 20 个 PcG 基因和至少 15 种不同的 trxG 基因。功能分析表明这些基因产生许多不同类型的蛋白质，而这些蛋白在真核生物中高度保守。PcG 基因的编码产物包括 DNA 结合蛋白（如 YY1）、组蛋白修饰酶（如 Ezh2）和其他具有 H3K27me3 亲和能力的 chromo 结构域的抑制性染色质相关因子（如 PC）。TrxG 基因编码转录因子（如 GAGA 或 Zeste）、ATP 依赖的染色质重塑酶（如 Brahma）和 HKMT，例如 Ash1 和 Trx（或其哺乳动物同源物 MLL、Set1 和 MLL 家族）。大多数情况下，trxG 和 PcG 家族的蛋白作为不同复合物中的组分而起作用，以建立稳定的染色质结构促进或沉默受发育调节的基因的表达（见第 11 和第 12 章）。

尽管最近取得了一些进展，但是包含 PcG 或 trxG 的复合物被靶定到发育调节的相关染色质区域的机制还不是很清楚。果蝇中，可遗传的基因抑制需要将 PcG 蛋白复合物募集到称为 Polycomb 应答元件（Polycomb response element，PRE）的 DNA 元件上。但在哺乳动物中尚未发现对应的 DNA 元件。尚不清楚 PcG 蛋白复合物是如何以一种 PRE 依赖的形式导致长距离基因沉默的，因为 PRE 通常与靶基因的转录起始位点相距数千个碱基。可以假想对 PcG 复合物的排斥或募集有可能是通过转录活性的改变或有效及无效 mRNA 加工加以区分的（Pirrotta 1998；Dellino et al. 2004；Schmitt et al. 2005）。现在的模型支持，PcG 结合是通过其与 DNA 结合蛋白的相互作用以及 PC 蛋白中 chromo 结构域与 H3K27me3 修饰的组蛋白的亲和力（Cao et al. 2002）。然而，PcG 复合物在体外能够与缺失组蛋白尾部的核小体相互作用（Francis et al. 2004），并且 PRE 元件具有较低的核小体密度（Schwartz et al. 2005）。对于这些矛盾现象最合理的解释是内源 PcG 的结合最初需要与 DNA 结合因子的相互作用，这种作用被与附近染色质区域的核小体和修饰的 H3K27me3 的相互作用而进一步稳定。显然，将有关 PcG 复合物是如何靶定到染色质区域，以及它们是如何介导沉默的现有证据联系起来还需要进行大量的研究。由于 PcG 复合物的差异性，这有可能是物种特异的（参见第 11 章）。

通常，TrxG 蛋白参与维持靶基因表达的活跃状态并克服（或阻止）PcG 介导的基因沉默。对这一转折过程的了解更少，但是最新证据表明一种基于 RNA 的机制能将 Ash1 募集到目标启动子上（Sanchez-Elsner et al. 2006）。一系列染色质结构的瞬时和稳定改变被认为是继发性的，可能由基因间转录所促使，因其能建立起一个开放的染色质区域并介导活跃的组蛋白置换。已报道的染色质改变包括 trxG HKMT（如 Trx 和

Ash1）催化的“激活性”的组蛋白-赖氨酸甲基化标签的掺入和这些标签的识别（如 WDR5 识别 H3K4me；Wysocka et al. 2005）。trxG ATP 依赖的染色质重塑因子例如 Brahma 的作用也是必需的，尽管还不完全清楚这些机制间是如何联系的（详见第 12 章）。

在正常细胞中，许多 PcG 和 trxG 蛋白协调作用，使抑制性的异染色质和活化的常染色质维持在一个受到严密调控的水平。在哺乳动物体细胞分裂间期的细胞核中，形态学分析显示着丝粒周边异染色质的组成型区域组成 15～20 个点状结构（图 3-16）。导致发育异常和癌症的细胞命运和增殖控制的失调常表现为异常的细胞核形态。例如，PML 白血病细胞（与 MLL 相关）的细胞核中失去了着丝粒周边异染色质点状结构（Di Croce 2005）。相反的，衰老（不增殖）的细胞则具有大的新出现的异染色质簇（Narita et al. 2003；Scaffidi et al. 2005）。因此，核形态学似乎是一种区分正常与异常细胞状态的很好的标记，这表明核结构可能在维持染色质的特定区域中具有调节作用。

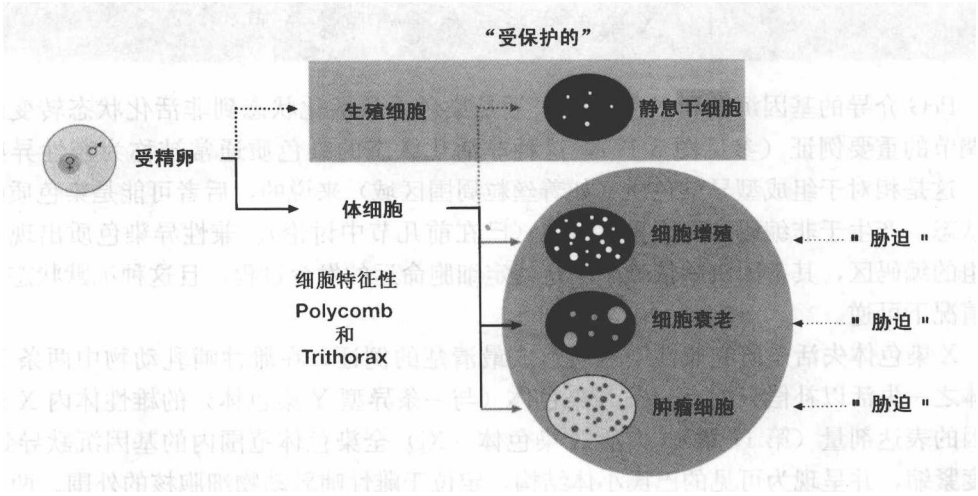



图 3-16　PcG 和 trxG 蛋白决定细胞特征性

在胚胎发生过程中，根据细胞分化能力的不同，建立起两种细胞组分：全能性的生殖细胞和只具有有限分化潜能的体细胞（包括干细胞）。生殖细胞或干细胞基因组表达能力的可塑性是通过抑制性组蛋白标记的减少体现的，着丝粒周边区域没有可见的抑制性组蛋白标记的点状结构。正常增殖细胞的典型核形态特征是存在 15～20 个异染色质点状结构。包含 Polycomb 和 Trithorax 的复合物在界定表观遗传特征性，并进一步界定不同世系的细胞特征性中共同发挥作用。它们也能够对于外界“胁迫”刺激做出应答，促进细胞增殖并进行适当的基因表达。衰老（老化）细胞中基因组可塑性和增殖能力丧失，其标志是异常大的异染色质区域和整体抑制性组蛋白标记的增加。然而，在高度增殖的肿瘤细胞中，存在着由于 PcG 和 TrxG 组蛋白修饰酶失调而导致的抑制性与活化性组蛋白标记平衡状态的改变。同时，这一现象伴随着核形态学的紊乱。

组蛋白修饰水平的高低是细胞正常与否的另一指标。这类变化多起因于 PcG（如 Ezh2）或 trxG（如 MLL）HKMT 的失调，从而导致肿瘤发生甚至转移（见第 15 节）。事实上，上述蛋白中任何一种的整体水平的上升都与前列腺癌、乳腺癌、多重骨髓瘤或白血病的发病几率呈正相关性（Lund and van Lohuizen 2004；Valk-Lingbeek et al.

2004)。在其他致瘤性转化的例子中，存在抑制性组蛋白标记的明显降低和乙酰化水平的整体升高（Seligson et al. 2005），它们导致了基因转录水平和基因组不稳定性的上升。显然，通过干扰组蛋白修饰酶而引发的染色质整体控制的改变影响了基因组的功能，并破坏了正常细胞基因正确的表达模式。

对于细胞的衰老，抑制性组蛋白标记的增加也是细胞功能发生紊乱的指标。这与相伴随的组蛋白乙酰化的减少，能够增强甚至增加沉默染色质的水平，降低细胞的可塑性，并驱使细胞进入一种抗增殖状态（Scaffidi et al. 2005）。这主要是一种与年龄相关的作用，尽管儿童早老症这种疾病能够使衰老提前。相反的，在缺失甲基转移酶 Suv39h 的突变体中，着丝粒周边区域抑制性的甲基化标记减少，细胞表现出更高程度的不朽性，不再发生衰老并伴有高度的基因组不稳定性（Braig et al. 2005）。这些例子表明，通常由 PcG 和 trxG 酶所介导的特征性组蛋白标记水平改变所代表的染色质异常与核形态的变化是疾病进程的一种重要指标。

## 13. X 染色体失活和兼性异染色质

PcG 介导的基因沉默和 X 染色体失活是染色质从活化状态到非活化状态转变的发育调节的重要例证（参见图 3-17），这种非活化状态的染色质通常被称为兼性异染色质。这是相对于组成型异染色质（如着丝粒周围区域）来说的，后者可能是染色质的默认状态，产生于非编码和高度重复区域（已在前几节中讨论）。兼性异染色质出现在基因组的编码区，其基因沉默依赖于决定特定细胞命运的发育过程，且这种沉默状态在某些情况下可逆。

X 染色体失活是研究兼性染色质形成最清楚的例证，在雌性哺乳动物中两条 X 染色体之一失活以补偿只拥有一条 X 染色体（与一条异型 Y 染色体）的雄性体内 X-连锁基因的表达剂量（第 17 章）。失活 X 染色体（Xi）全染色体范围内的基因沉默导致 Xi 高度紧缩，并呈现为可见的巴氏小体结构，定位于雌性哺乳动物细胞核的外围。两个 X 染色体是如何被识别，又如何选定一条特定的 X 染色体并使之失活，这些都是当今表观遗传研究面临的挑战性问题。

X 染色体失活需要一个很大的（约 17 kb）非编码 RNA，*Xist*，它似乎是 Xi 染色质重塑的初始诱因。尽管 *Xist* 与其反义转录物 *Tsix*（只在 X 染色体失活之前表达）之间有形成 dsRNA 的潜能，但还没有 RNAi 机制导致 X 染色体失活的有力证据。X 染色体失活中心（X inactivation center，XIC）可能作为 *Xist* RNA 在 DNA 上的“入口”或“停泊”位点（推测为 X 染色体富集的特定重复 DNA 元件），在 *Xist* 的结合并作为骨架分子以顺式方式装饰 Xi 的过程中发挥作用。*Xist* 促进 PRC1 和 PRC2 复合物的募集和发挥功能，后者参与建立稳定的失活 X 染色体。例如，PRC2 复合物包含 HKMT 染色质修饰酶-EZH2，催化 H3K27 的三甲基化。PRC1 复合物结合则同时被 H3K27me3 和组蛋白修饰非依赖方式所促进，而复合物中的其他组分如 Ring1 蛋白能够将 H2A 泛素化。这就是 PcG 复合物的异质性，也就是说不同组分能够独立于其他组分而发挥作用。染色质修饰、PcG 复合物的结合以及后来一系列组蛋白变体沿 X 染色体的掺入和

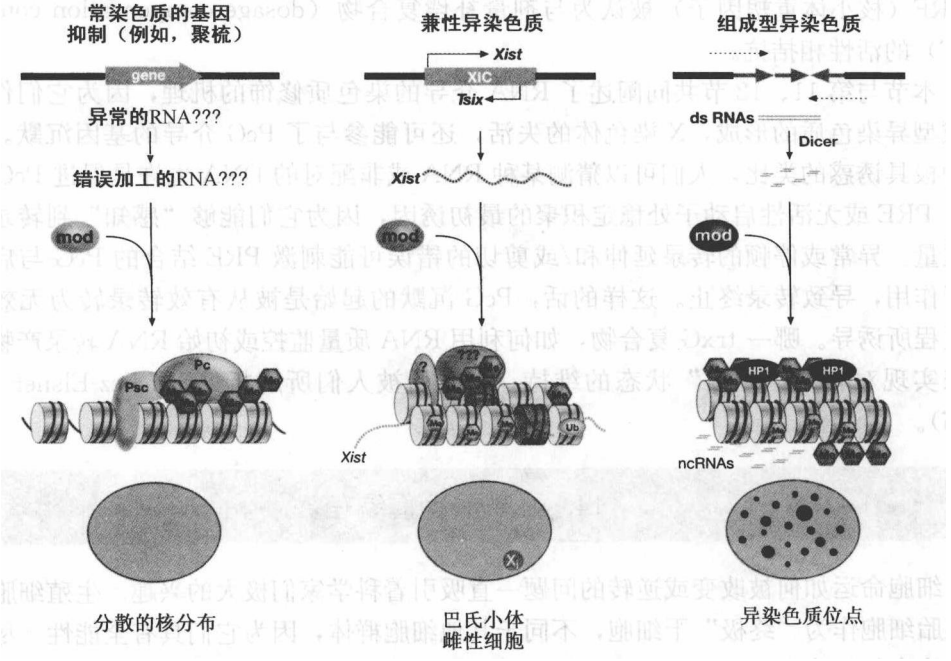



图 3-17　RNA 介导的染色质抑制状态的诱导

不同形式的染色质沉默具有不同的初始信号，但许多都可能与 RNA 转录物相关（从异常转录物，到 *Xist* RNA，再到 dsRNA），取决于所作用的 DNA 序列。它引发了一系列的染色质变化，包括组蛋白修饰的组合（H3K9、H3K27 和 H4K20 甲基化），抑制性蛋白或复合物（如 Pc 或 HP1）的结合，DNA 甲基化和组蛋白变体（如失活 X 染色体上的 macroH2A）的存在。兼性或组成型异染色质在细胞核内均呈现可见的聚集状结构。常染色质抑制则不能通过细胞核的形态学模式来确定。

大量的 DNA 甲基化都参与了整个失活 X 染色体上兼性异染色质结构的形成。一旦一个稳定的异染色质结构已经形成，其维持不再需要 *Xist* RNA 的存在（Avner and Heard 2001；Heard 2005）。基因组印记是另一种类似的单等位基因沉默形式，它也通过使用非编码 RNA 或反义 RNA，以亲本特异性的形式来沉默等位基因中的一个拷贝（第 19 章）。现在还不清楚 *Dicer* 突变的小鼠胚胎干细胞是否及如何影响 X 染色体失活或基因组印记的。

作为经典的表观遗传调控机制的通用范例，剂量补偿在其他模式生物，尤其是 *C. elegans*（Meyer et al. 2004；第 15 章）和 *Drosophila*（Gilfillan et al. 2004；第 16 章）中已被阐述。然而，目前还不清楚这种剂量补偿是否存在于鸟类中，尽管它们是异配子生物。在果蝇中，性别之间的剂量补偿并不是通过雌性的 X 染色体失活，而是通过雄性体内单一 X 染色质上基因表达的两倍上调来实现的。有趣的是，两种非编码 RNA，*roX1* 和 *roX2*，被认为是非常重要的元件，它们的表达具有雄性特异性。虽然类似的机制细节可能存在于果蝇和哺乳动物间，但显然，活性的染色质重塑和组蛋白修饰，尤其是 MOF 依赖的雄性 X 染色体的 H4K16 乙酰化，在果蝇的剂量补偿中起了重要的作用。组蛋白修饰活性，例如 MOF 组蛋白乙酰转移酶，是怎样被精确地定位于雄性 X 染色体

上仍然是未来研究所面临的一个挑战。此外，ATP 依赖的染色质重塑活性，例如，NURF（核小体重塑因子）被认为与剂量补偿复合物（dosage compensation complex，DCC）的活性相拮抗。

本节与第 11、12 节共同阐述了 RNA 介导的染色质修饰的机理，因为它们作用于组成型异染色质的形成，X 染色体的失活，还可能参与了 PcG 介导的基因沉默。鉴于这种极具诱惑的类比，人们可以猜测某种 RNA 或非配对的 DNA 也许是促进 PcG 复合物在 PRE 或无活性启动子处稳定积聚的最初诱因，因为它们能够“感知”到转录进程的质量。异常或停顿的转录延伸和/或剪切的错误可能刺激 PRE 结合的 PcG 与启动子相互作用，导致转录终止。这样的话，PcG 沉默的起始是被从有效转录转为无效转录的过程所诱导。哪一 trxG 复合物，如何利用 RNA 质量监控或初始 RNA 转录产物的加工来实现对转录“开启”状态的维持，正开始被人们所解析（Sanchez-Elsner et al. 2006）。

## 14. 细胞命运重编程

细胞命运如何被改变或逆转的问题一直吸引着科学家们极大的兴趣。生殖细胞和早期胚胎细胞作为“终极”干细胞，不同于其他细胞群体，因为它们具有全能性。尽管哺乳动物中的细胞命运决定导致约 200 种不同的细胞类型存在，本质上，只存在两种主要的分化转变：从干细胞（多能性的）到完全分化的细胞和由静息细胞（静止或处于 $G_0$ 期）到增殖细胞。这代表了大量处于过渡状态细胞的极端情况，这与哺乳动物发育过程中表观基因组“组分”的多样性是一致的。胚胎发生过程中，在从受精卵母细胞到胚泡，再到着床、原肠胚形成、器官发育和胚胎发育的过程中，表观遗传修饰存在动态性的增加。大部分的修饰或印记能够通过将分化细胞的细胞核移植到去除细胞核的卵母细胞的细胞质中得以去除。然而，一些标签仍然存在，并限制克隆胚胎的正常发育，一些甚至能够以生殖细胞修饰（g-mod）的形式被遗传（参见图 3-18），在哺乳动物细胞中，这可能包括 DNA 甲基化。

尽管大部分组织不能够被重编程，肝脏再生或肌细胞修复是哺乳动物组织中的特例，它们能够在遭到损伤之后再生。在其他生物体中，例如植物和蝾螈，特定的体细胞能够对其表观基因组进行重编程并重新进入细胞周期来重建丢失的或损伤的组织（Tanaka 2003）。一般说来，体细胞的重编程是不可能的，除非利用去核卵母细胞的核移植技术（nuclear transfer，NT）来“操纵”早期发育的重演。这最早在克隆爪蟾中，最近又在第一只克隆哺乳动物“多莉”的产生（Campbell et al. 1996，见第 22 章）中被论证。

哺乳动物体细胞重编程的高效性面临三种已知的主要障碍。第一，特定体细胞的表观遗传标签（如抑制性的 H3K9me3）能够通过体细胞分裂进行稳定传播，并抵制卵母细胞的重编程。第二，体细胞的细胞核不能够重演受精胚胎中所发生的重编程的不对称性，因为后者从父本和母本单倍体基因组继承了不同的表观遗传标记（见 Mayer et al. 2000；van der Heijden et al. 2005；第 20 章）。第三，对胚胎及胎盘发育非常重要的印

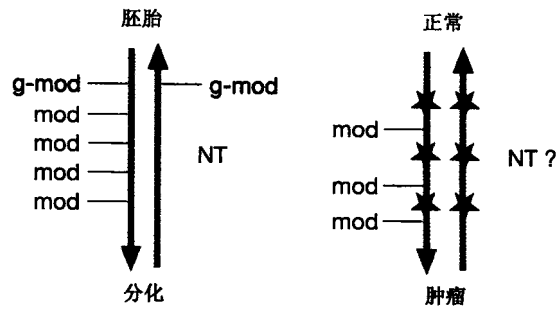



图 3-18　核转移导致的重编程

在一个个体的生命周期中，表观遗传修饰（mod）在不同的细胞谱系中是必需的（左）。体细胞核移植（nuclear transfer，NT）逆转了终端分化的进程，去除了绝大多数的表观遗传标记。然而一些在生殖细胞中也存在的修饰（g-mod）不能被抹除。在一系列遗传突变（星号）所导致的致瘤性转化（从正常细胞到肿瘤细胞）中，表观遗传损伤（mod）逐渐累积。它们能通过核移植的重编程被去除，而突变则不能。这说明遗传学与表观遗传学在肿瘤发生中相互影响。图片由 R. Jaenisch 提供。

记化区域的传承不能在核转移实验中被如实保持（Morgan et al. 2005）。大部分克隆胚胎的夭折表明混乱的表观遗传印记是正常发育的瓶颈，这也可能是辅助生殖技术（assisted reproductive technologies，ART）效率低和克隆动物活力弱的原因。

相对于体细胞而言，胚胎干细胞（embryonic stem cells，ES）具有更强的重编程潜能。静息细胞（干细胞常有的一种特性）中整体 H3K9me3 和 H4K20me3 水平的降低有可能意味着其表观基因组具较强的可塑性（Baxter et al. 2004）。这与“不朽的”单细胞生物体（如酵母）拥有一个大致“开启”和活化的基因组，而缺少一些抑制性表观遗传机制的现象相一致。

受精后，哺乳动物正常表观遗传重编程的另一个特性是其明显的不对称性。这首先应归结于雄性及雌性生殖细胞中不同的表观遗传特性（见第 19 章和第 20 章）。精子基因组大部分由精蛋白组成，虽然还具有残留的，但仍具重要意义的 CENP-A（一种 H3 组蛋白变体）和其他可能的表观遗传印记（Kimmins and Sassone-Corsi 2005），而卵母细胞则具有常规的包含核小体的染色质。受精后，精细胞与卵母细胞的单倍体基因组会经历新一轮的重编程，包括 DNA 去甲基化和组蛋白变体的替换。这些修饰能够在核融合之前的第一轮细胞周期中提高或平衡两套亲本基因组之间的表观差异。在胚胎（即内细胞群或 ICM）和胚外［即滋养外胚层（TE）和胎盘］组织的分化过程中，不同世系中建立起不同的 DNA 甲基化和组蛋白修饰模式（Morgan et al. 2005）。体细胞克隆不能够如实地重演这些重编程模式，表现为第一个细胞周期中的快速但不全面的 DNA 去甲基化，以及 ICM 和 TE 细胞中 DNA 甲基化和组蛋白赖氨酸甲基化的异常。

印记基因的命运是与体细胞重编程紧密相关的问题。对于正常的胚胎发育，印记化区域正确的等位基因表达是必需的（见第 19 章）。单亲胚胎生殖实验对此做出了论证（McGrath and Solter 1984；Barton et al. 1984；Surani et al. 1984）。雄源胚胎（两套基因组都来源于雄性）的胚胎发育较为迟缓，但胚外组织（如胎盘）却出现过度增殖。

雌源或单性生殖的胚胎中（两套基因组都来源于雌性），胎盘发育不全。因此，在生殖细胞中，亲本特异性印记必须在擦除原有标签之后重新建立（见第 20 章）。科学家们相信，这发生在大约 100 个或更多的与胚胎及胎盘发育资源供给系统相关的印记化基因（如 *Igf2* 生长因子）上。有趣的是，有证据显示，通过辅助生殖技术（assisted reproductive techniques，ART）或核移植得到胚胎在体外培养过程中，印记可能被打乱（Maher 2005）。

## 15. 癌

细胞的自我更新与分化之间存在着一个微妙的平衡。致瘤性转化（也常被称为肿瘤发生）是细胞发生一系列改变的过程，包括细胞增殖的失控、检验点（checkpoint）控制机制的丢失导致染色体异常和基因组非整倍性的累积，以及细胞分化的错误调节等（Lengauer et al. 1998）。这通常被认为起因于至少一个遗传损伤，如点突变、缺失或者易位破坏了抑癌基因（tumor suppressor gene）或癌基因（oncogene）（Hanahan and Weinberg 2000）。可能是癌细胞中的抑癌基因被沉默，也可能是正常基因（原癌基因）通过显性突变或过量表达被激活成癌基因。重要的是，异常表观遗传修饰的累积也伴随着肿瘤细胞的形成（见第 24 章）。这些表观遗传的改变包括 DNA 甲基化模式、组蛋白修饰和染色体结构的改变（图 3-19）。因此，致瘤性转化是一个复杂的多步骤过程，它通过遗传或表观遗传事件导致随机性癌基因激活和/或抑癌基因沉默，也即所谓的“Knudson 双重打击”理论（Feinberg 2004；Feinberg and Tyko 2004）。举例来讲，视网膜母细胞瘤基因（retinoblastoma，Rb，一种抑癌基因）的沉默导致检验点的消失，不但有利于（癌）细胞的增殖，还通过影响与维持基因组稳定性的染色体结构相关的下游功能来促使肿瘤发生的“二次打击”（Gonzalo and Blasco 2005）。癌基因（如 *myc* 基因）的不恰当激活也会导致类似的后果。

当前研究提出的问题之一是，异常的表观遗传变化对肿瘤发生率及其整个进程究竟有多大程度的作用？以黑色素瘤细胞核为供体的核转移（nuclear transfer，NT）实验回答了这一问题（Hochedlinger et al. 2004）。尽管供体细胞的每一个遗传损伤均被保留，但核转移将抹去其表观遗传特性。随后对克隆鼠肿瘤发生率的研究表明，新产生的肿瘤谱变化非常大，这符合不同组织中表观遗传修饰对刺激肿瘤发生有着不同作用的观点。

DNA 低甲基化（hypomethylation）（与之相对的是高甲基化，hypermethylation）可以发生在某些特定区域或整个染色体区都可能发生。事实上，DNA 低甲基化是第一种与癌症相联系的表观遗传转换类型（Feinberg and Vogelstein 1983）。它是肿瘤细胞一种广为存在的表型。在单基因水平上，DNA 低甲基化可以通过对原癌基因的激活，对造成异常细胞功能的基因的去抑制，或者印记基因的双等位基因表达（也称基因印记丢失，LOI），进而导致肿瘤发生（见第 23 章、24 章）。在更广的基因组水平上，大量的 DNA 低甲基化，特别是组成型异染色质区的 DNA 低甲基化使细胞易于发生染色体易位和染色体非整倍化，从而导致肿瘤发生。这种效果在 Dnmt1 突变体中得以体现

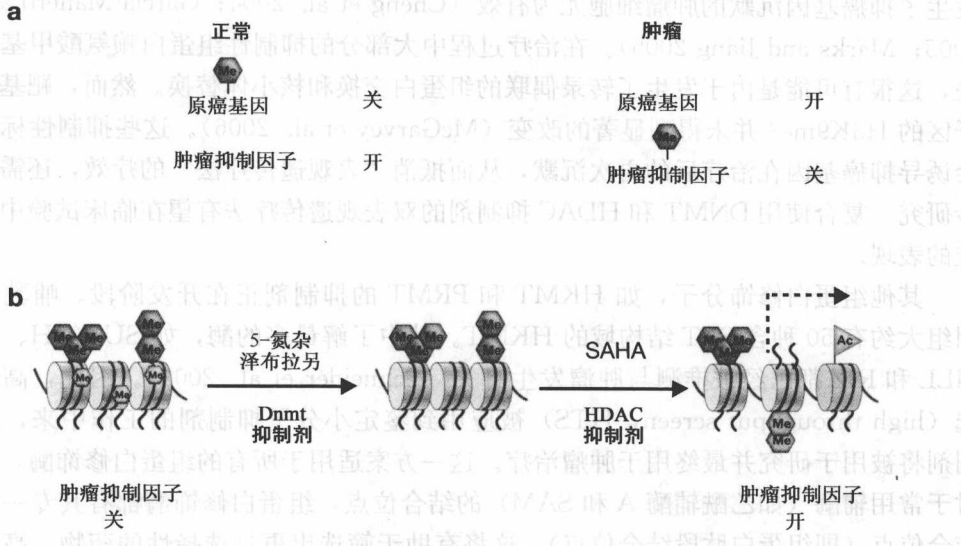



图 3-19　肿瘤中的表观遗传修饰

(a) 肿瘤中致癌位点的异常表观遗传修饰一般导致癌基因的去抑制或抑癌基因的沉默。已知能改变正常细胞的表观遗传标记包括 DNA 甲基化、抑制性组蛋白甲基化和组蛋白去乙酰化。(b) 以一个抑癌基因座为例，示意表观遗传治疗药物作用于癌细胞，使其染色质模版发生改变。Dnmt 抑制剂导致 DNA 甲基化丢失，HDAC 抑制剂导致组蛋白乙酰化标记的掺入，继而引发下游修饰，包括激活性组蛋白甲基化标记和组蛋白变体的掺入。这些染色质改变共同导致基因的重新表达。SAHA：辛二酰苯胺异羟肟酸。

(Chen et al. 1998)。DNA 低甲基化所引起的基因组不稳定性可能是由转座子活化产生的诱变作用所导致的。随着人们对抑制性组蛋白修饰在着丝粒和端粒处维持异染色质的重要作用的认识，越来越多的证据显示这些标记的丢失，也会引起基因组不稳定性，从而导致肿瘤发生 (Gonzalo and Blasco 2005)。

相反地，DNA 高甲基化集中出现在多种肿瘤中具有 CpG 岛的启动子区。这种异常的 DNA 高甲基化所导致的抑癌基因沉默对癌的恶化尤其重要。近来的研究揭示染色质修饰与 DNA 甲基化之间存在众多联系，抑癌基因的沉默，往往有多种表观遗传机制的参与。举例来讲，抑癌基因 *p16* 或 *hMLH1* 的沉默是通过 DNA 甲基化和抑制性组蛋白赖氨酸甲基化共同实现的 (McGarvey et al. 2006)。

很多类型的肿瘤都显示出染色质修饰分子的异常。某些组蛋白修饰蛋白变的具有成瘤性，例如 PcG 蛋白 EZH2 和 trxG 蛋白 MLL，它们通过干扰细胞的表观遗传特征性来实现其影响，最终导致不恰当的基因转录沉默或激活 (Valk-Lingbeek et al. 2004; Schneider et al. 2002)。很显然，表观遗传特征性对细胞的功能是至关重要的。事实上，正如前列腺癌发生的研究结果所展示的那样，特征性的组蛋白乙酰化和甲基化标记模式被证明是某些肿瘤发生的标志 (Seligson et al. 2005)。

靶向性抑制染色质修饰效应酶的药物研发开创了肿瘤治疗的新篇章 (图 3-19b)。DNMT 和 HDAC 抑制剂的使用处于新一代肿瘤疗法的临床试验最前沿。Zebularine 和 SAHA 分别是上述两种酶的抑制剂 (图 3-19)。由于这种疗法具转录激活作用，它们对

发生了抑癌基因沉默的肿瘤细胞尤为有效（Cheng et al. 2004；Garcia-Manero and Issa 2005；Marks and Jiang 2005）。在治疗过程中大部分的抑制性组蛋白赖氨酸甲基化被丢失，这很有可能是由于发生了转录偶联的组蛋白交换和核小体替换。然而，靶基因启动子区的 H3K9me3 并未得到显著的改变（McGarvey et al. 2006）。这些抑制性标记是否会诱导抑癌基因在治疗后的再次沉默，从而抵消“表观遗传疗法”的疗效，还需要进一步研究。复合使用 DNMT 和 HDAC 抑制剂的双表观遗传疗法有望在临床试验中取得更佳的表现。

其他组蛋白修饰分子，如 HKMT 和 PRMT 的抑制剂正在开发阶段。哺乳动物基因组大约有 50 种含 SET 结构域的 HKMT。其中了解最多的酶，如 SUV39H、EZH2、MLL 和 RIZ 都已经被推测与肿瘤发生有关（Schneider et al. 2002）。因此，高通量筛选（high throughput screen，HTS）被应用到鉴定小分子抑制剂的工作中来，这些抑制剂将被用于研究并最终用于肿瘤治疗。这一方案适用于所有的组蛋白修饰酶，因为相对于常用辅酶（如乙酰辅酶 A 和 SAM）的结合位点，组蛋白修饰酶都有其专一性底物结合位点（即组蛋白肽段结合位点），这将有助于筛选出更具选择性的药物。高通量筛选已经成功地针对 HDAC（Su et al. 2000）、PRMT（Cheng et al. 2004）和 HKMT（Greiner et al. 2005）进行了筛选。

为了实现从基础研究向应用研究的转变，任何一种组蛋白修饰酶抑制剂的功效和用途都需要经历理论模型以及实际试验两种手段的考验。例如，针对 MLL 或 EZH2 的专一性抑制剂可能是白血病或前列腺癌的有价值的治疗药物。而 SUV39H 抑制剂虽则直觉上不像有用的药物，因为 SUV39H 对于维持组成型异染色质和基因组稳定性是必须的，然而仍有可能优先对癌细胞产生影响。

对 HDAC 抑制剂 SAHA 的分析显示，它还可能通过转录重激活以外的其他途径发挥作用（Marks and Jiang 2005）。例如，HDAC 抑制剂能够使细胞对染色质损伤变得敏感，抑制有效 DNA 修复和允许基因组不稳定性从而引发肿瘤细胞的凋亡。

评价双联疗法的疗效时仍必须监测以上观察到的现象。然而，根据目前已有的结果，可以预见 HDAC 和 HKMT 抑制剂的联合疗法能更具选择性的通过诱导信息紊乱和染色质灾变来杀死肿瘤前体细胞。值得期待的是，今后的研究将会为表观遗传肿瘤疗法鉴定出可行的备选药物。

## 16. 表观遗传控制的意义

哺乳动物基因组编码的蛋白质中约有 10%被用于转录和染色质调节（Swiss-Prot database）。由于哺乳动物基因组中含有 $3\times10^9$ 个碱基对，它必然有约 $1\times10^7$ 个核小体。这提供了大量的潜在调节信息，包括 DNA 结合作用、组蛋白修饰、组蛋白变体、核小体重塑、DNA 甲基化和非编码 RNA。然而，转录调节过程本身相当复杂，经常需要组装巨大的多蛋白质复合体（超过 100 种蛋白质）来保证对某一选定启动子的起始、延伸和对信使 RNA 正确的加工。如果 DNA 序列特异性调节都非常精细，那么我们可以预见，对 DNA—组蛋白聚合体的动态低亲和力相互作用将会需要更加精细的调节。有鉴

于此，与一种表观遗传状态相对应的修饰往往不止一种。就像实验证据所表明的那样，更多的是在某染色质区域内，由几个（乃至更多）信号的复合效应或累积作用来共同稳定和传播其表观遗传状态（Fischle and Allis 2003b；Lachner et al. 2003；Henikoff 2005）。

大多数情况下，转录因子的结合是短暂的，并在随后的细胞分裂过程中丢失。对于持续性基因表达模式，在后续的每次细胞分裂过程中都需要转录因子。因此，表观遗传控制能够以染色质模板为媒介，通过信息的遗传传递来增强原始信号（如启动子刺激、基因沉默、着丝粒定义）向后代（但非永久性）细胞的传递（图 3-20）。有趣的是，*S. pombe* 中依赖 Swi6 的表观遗传花斑能够被组蛋白修饰（多为 H3K9me2）在多次细胞分裂（包括有丝分裂和减数分裂）中抑制（Grewal and Klar 1996）。人们利用果蝇做了相似的研究，通过转录激活因子的单次刺激，实现 *Hox* 基因表达在雌性生殖细胞系中的细胞记忆传递（Cavalli and Paro 1999）。在这两例中，表观遗传记忆都是通过染色质变化所介导的，包括了不同的组蛋白修饰，很可能还有组蛋白变体的掺入。

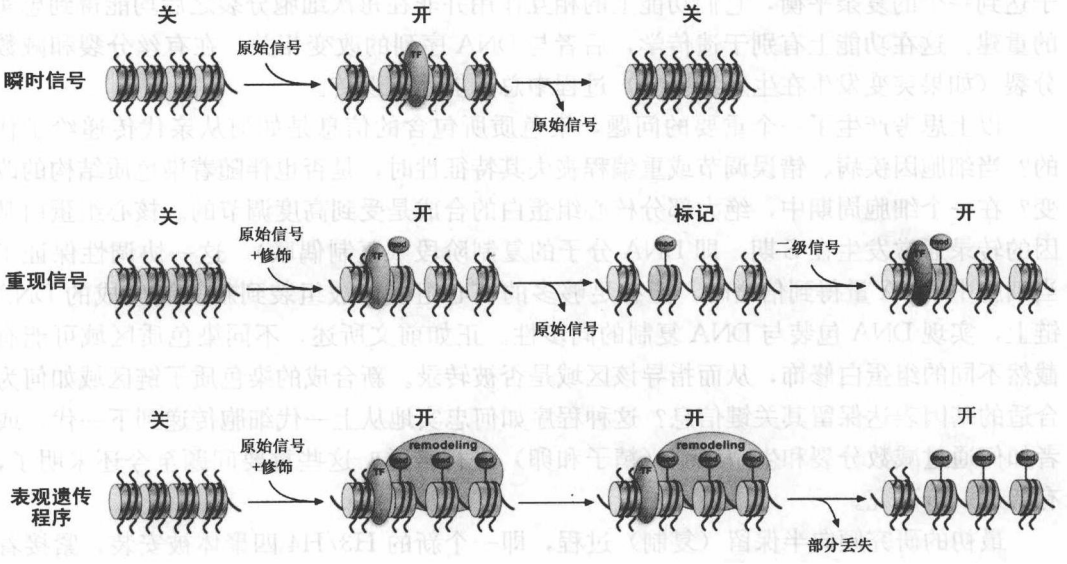



图 3-20　初始信号的表观遗传增强（记忆/遗传）

经典遗传学认为基因表达依赖于转录因子（transcription factor，TF）的存在与结合。这种因子（即初始信号）的去除导致基因表达的消失，因此属于瞬时激活信号。染色体结构与基因表达相关，一些结构是抑制性的，而一些是激活的。一个位点的激活有可能通过初始信号发生，从而导致下游染色质结构改变，包括组蛋白的激活性共价修饰和组蛋白变体（如 H3.3）对核心组蛋白的替换。通过细胞分裂，这种染色体结构可能只在激活信号（指的是“重现信号”）存在的情况下才能被重建。表观遗传记忆使染色质在细胞分裂过程中，即使是在最初激活信号消失的情况下，仍然维持其状态。这种记忆系统并不是绝对的，它包括对重塑染色质结构的多层次表观遗传调节。染色质的动态特性表明，即便某种染色质状态可能在有丝分裂过程中是稳定的，它仍倾向于发生改变，从而影响表观遗传记忆的时间。

如果组蛋白修饰共同作用，可能会在染色质模板上留下印记以帮助标识该核小体，特别是如果这一信号能随 DNA 复制而复制（图 3-20）。为了更加稳定的继承，组蛋白

修饰、组蛋白变体掺入和染色质重塑的相互协作将使某片染色质区域转变为持续的结构，并在多次细胞分裂中得以传递。尽管列举的是转录“开启”状态的遗传性，类似的协作也发生在抑制性表观遗传机制中，它们将染色质区域锁入一种更加稳定的沉默状态，这种沉默状态被额外的 DNA 甲基化所加强。

因此，DNA 双螺旋可以被认为是一种自组织多聚体，能通过其在染色质上的排列实现对表观遗传控制的响应，并将初级信号放大成为一种更为长期的“记忆”。此外，许多组蛋白修饰很可能是为了响应细胞内在或外来刺激所产生的。与此一致的是，染色质修饰酶需要辅酶如 ATP（激酶）、乙酰辅酶 A（HAT）和 SAM（HKMT）等，它们的细胞含量受环境变化（如饮食）的影响。因此，环境变化能导致一种更为动态的或稳定的 DNA-组蛋白多聚体。NAD-依赖的 HDAC、Sir2 就是很好的例子，它们是营养物质和生命周期/衰老细胞的“感应器”（Guarente and Picard 2005；Rine 2005）。解释这些环境影响是怎样成为生物学上的表观遗传信号，且这些信号是如何被解读、翻译和遗传是目前表观遗传研究的核心内容。然而值得强调的是，表观遗传控制需要众多影响因子达到一个的复杂平衡，它们功能上的相互作用并非在每次细胞分裂之后均能得到忠实的重建。这在功能上有别于遗传学，后者与 DNA 序列的改变相关，在有丝分裂和减数分裂（如果突变发生在生殖细胞中）过程中总是能稳定遗传。

以上思考产生了一个重要的问题，染色质所包含的信息是如何从亲代传递给子代的？当细胞因疾病、错误调节或重编程丧失其特征性时，是否也伴随着染色质结构的改变？在一个细胞周期中，绝大部分核心组蛋白的合成是受到高度调节的。核心组蛋白基因的转录通常发生在 S 期，即 DNA 分子的复制阶段（复制偶联）。这一协调性保证了当细胞中 DNA 量得到倍增时，能有足够多的核心组蛋白被组装到新复制生成的 DNA 链上，实现 DNA 包装与 DNA 复制的同步性。正如前文所述，不同染色质区域可能有截然不同的组蛋白修饰，从而指导该区域是否被转录。新合成的染色质子链区域如何为合适的基因表达保留其关键信息？这种程序如何忠实地从上一代细胞传递到下一代，或者如何通过减数分裂和生殖细胞（精子和卵）进行传递？这些重要问题至今还未明了，有待进一步研究。

最初的研究倾向半保留（复制）过程，即一个新的 H3/H4 四聚体被安装，紧接着两个新的 H2A/H2B 二聚体再结合上去，但最新的实验结果对上述假说提出了挑战。在新的研究模型中，“新的”可能已携带某些转录后修饰的 H3 和 H4 多肽以新合成的 H3/H4 组蛋白二聚体形式，与伴随着 DNA 链分离的“旧的”H3/H4 二聚体共同组装成一个核小体。如果事实如此的话，那么修饰的亲代 H3/H4 二聚体会和新合成的二聚体出现在同一条 DNA 链上。这种共存可能进而指导对新加入的二聚体进行适当的修饰（Tagami et al. 2004）。这个模型极具吸引力，它可能有助于解释组蛋白修饰的继承，继而解答 DNA 复制和细胞分裂过程中表观遗传信息的传递等难题。然而，要想解答染色质标记在细胞分裂中的传递，还需要更多的证据证明这个模型或其他模型的正确性。

在本章的结语部分，尚有一个疑问是表观遗传控制是否与基础遗传学原理有着根本性的差异？尽管我们期望将 Waddington 的表观遗传全貌看做是一片片激活性或抑制性组蛋白修饰区域坐落于连续的染色质聚合体上，但是这一看法易被过度解读。仅仅在最

近几年，我们才了解了一些可能与组蛋白修饰传递有关的主要酶系统。这形成了目前我们关于一些组蛋白标记稳定性及其可继承性的观点。另外，值得强调的是最新的研究表明在发育异常或肿瘤形成中常有染色质修饰活性的突变，例如核小体重塑（Cho et al. 2004；Mohrmann and Verrijzer 2005）、DNMT（Robertson 2005）、HDAC 或者 HMKT 等（Schneider et al. 2002）。这是遗传控制具有最终决定权的实例。因此，这些突变小鼠体内的肿瘤发生往往被认为是一种“遗传”疾病。相反地，不因基因突变而发生的核小体结构、DNA 甲基化和组蛋白修饰的改变才是“真正的”表观遗传异常。这些更“可塑系统”的优秀例证包括早期胚胎发育中的随机决定、核移植引起的重编程、转录记忆、基因组印记、X 染色体失活的嵌合性、着丝粒特征性和肿瘤进程等。因此，遗传与表观遗传是紧密相关的现象，细胞分裂后的可继承性是它们固有的特性。对遗传控制来说如果演变发生在生殖细胞中，生殖细胞功能也会受影响。对于其他一些，常常被轻易归类为“表观遗传”的修饰，我们不知道它们是否仅仅反应了一个针对外界环境改变的次要的、短暂的响应，还是能在数代（但非永久）的细胞分裂过程中被保留下来的具有深远意义的表型改变，甚至偶尔这种改变还会对生殖细胞发生影响。虽然今天我们拥有大量新的表观遗传机制的知识，但很少或者毫无新的证据来支持拉马克学说（相对于达尔文进化论的拉马克用进废退学说，译者按）。

## 17. 表观遗传研究所面临的重大问题

本书讨论表观遗传现象发生的基本概念和一般原理，正如这些现象本身一样令人迷惑。我们的最终目标是从丰富的生物学背景中，勾画出对指挥和塑造这些概念的机制的最新理解，并将其呈现给读者。短短几年中，表观遗传学研究涌现出大量令人瞩目的观点和突破性进展，然而一些长期存在的疑问依然没有答案（图 3-21）。依据这些进展做出概括性结论并提出统一的规则对我们极具诱惑，然而由于担心还会有许多例外会打破这些规则，我们谨慎地避免这一倾向性。例如，物种之间明显存在着显著的差异性。值得一提的是，从单细胞到多细胞生物，组蛋白修饰、组蛋白变体、DNA 甲基化的程度和类型以及 RNA 干扰机制在不同物种中均具有显著差异。

然而，我们有足够多的理由将大量的精力投入到探索表观遗传现象分子机制的研究中去。精心设计的生物化学和遗传学研究已经在这些通路的许多功能性研究中取得了空前的成功。因此，可以预见到，对不同细胞类型的表观遗传转变（例如，干细胞与分化细胞；静息细胞与增殖细胞）进行细致的分析将会揭示多能性（pluripotency）的标志性特征（Bernstein et al. 2006；Boyer et al. 2006；Lee et al. 2006）。通过与疾病状态和肿瘤发生进行比较，将很可能有助于判断出在（细胞）正常的分化过程中哪些染色质变化具有重要意义。例如，对正常细胞、肿瘤细胞或胚胎干细胞进行大规模、全染色体“表观遗传全貌”测定的研究（Brachen et al. 2006b；Squazzo et al. 2006；Epigenomics AG，ENCODE，GEN-AU，*EPIGENOME* NoE），预期将会产生新的知识，带来创新性疾病治疗手段并催生一个国际大联盟来共同实现整个人类表观基因组的绘制（Jones and Martienssen 2005）。可以想象，不同组蛋白修饰水平的物种差异性，例如酿酒酵母

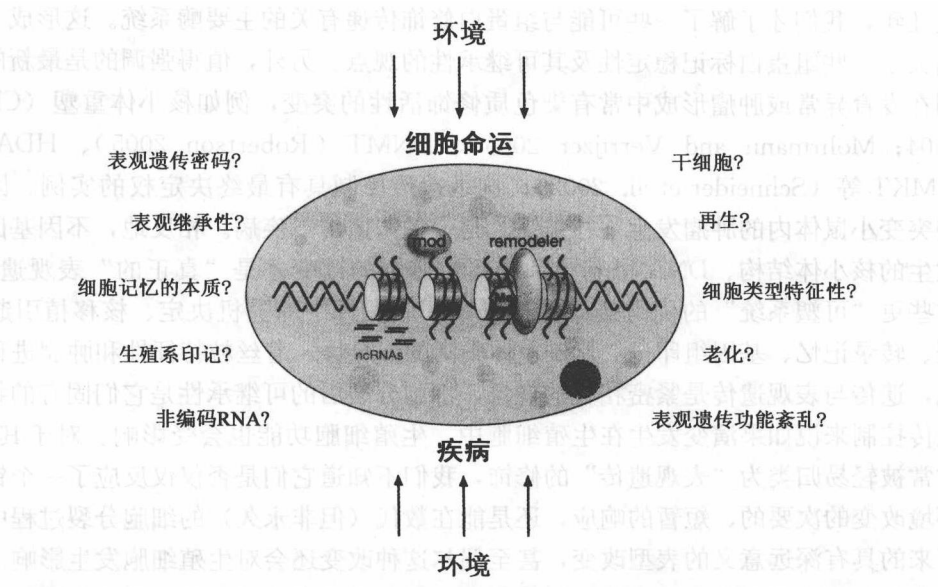



图 3-21 表观遗传学研究领域的重大问题

表观遗传研究所使用的众多实验系统揭示了关于表观遗传调节机制的许多新观点。有很多疑问仍然存在，有待于新的和已有的实验模式系统进一步证实。

和拟南芥中抑制性的组蛋白赖氨酸三甲基化水平明显偏低，很可能反映了这些生物强大的增殖和再生潜能，而多细胞生物系统则具有更受约束的发育程序。此外，RNA 干扰系统、组蛋白赖氨酸甲基化和 DNA 甲基化功能上的相互联系将继续为研究发育过程中细胞命运决定的复杂机制带来更多令人兴奋的惊喜。相似的，对核小体重塑体系的动力学和专一性的深刻理解也将对此作出贡献。我们预言，更多“奇异”的，通过修饰组蛋白和非组蛋白底物来催化表观遗传转变的酶活性将会被发现。上述机制诱导的染色质改变在很大程度上看上去是在扮演着一个环境信号过滤器的角色。因此，这些知识最终有可能通过对患者体内导致衰老、疾病和肿瘤的表观遗传特性进行重新设置，并用于改进治疗策略。这包括组织再生、治疗性克隆（利用 ES 细胞及其衍生物）和成体干细胞治疗。相信这些策略将会延长细胞的生命周期，调节细胞对外界刺激的抗逆应答，逆转疾病进程和提高生殖辅助技术。我们预言，对细胞多能性和全能性染色质基础的理解，将会处于对干细胞生物学及其潜在的医疗应用的核心地位。

许多基本的表观遗传学问题仍没有解决。例如，在相同的细胞核环境里，对于携带相同 DNA 序列的两条染色体，是什么将其中一条与它的等位基因相区分？表观遗传信息是如何被继承并传播的？细胞记忆的分子机制是什么？生殖细胞中存在使其基因组维持一种全能状态的某种表观遗传印记吗？我们期待下一代的研究人员拥有老一辈遗传与表观遗传学研究者的信心和热情，拿出足够的勇气来应对这些问题。

总之，Mendel 描述的遗传学原理很可能主导了绝大多数发育和外在表形。然而，这些法则有例外，有时会揭示出一些过去被忽视或不为人知的指导遗传的新原理和新机制。这本书旨在将对与 DNA 突变无关的表形变化的最新的理解展现给读者。我们希望

这本书描述的系统和概念能够给新时期的学生和研究人员，以及那些对表观遗传现象好奇并被吸引的人们提供一个有用的基础。

朱 冰 孙方霖 译校

## 参 考 文 献


Ahmad K. and Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9:** 1191–1200.

Allfrey V.G., Faulkner R., and Mirsky A.E. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci.* **51:** 786–794.

Almeida R. and Allshire R.C. 2005. RNA silencing and genome regulation. *Trends Cell Biol.* **15:** 251–258.

Aparicio O.M., Billington B.L., and Gottschling D.E. 1991. Modifiers of position effect are shared between telomeric and silent mating-type loci in *S. cerevisiae*. *Cell* **66:** 1279–1287.

Avery O.T., Macleod C.M., and McCarty M. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus Type III. *J. Exp. Med.* **79:** 137–158.

Avner P. and Heard E. 2001. X-chromosome inactivation: Counting, choice and initiation. *Nat. Rev. Genet.* **2:** 59–67.

Bannister A.J. and Kouzarides T. 2005. Reversing histone methylation. *Nature* **436:** 1103–1106.

Bannister A.J., Zegerman P., Partridge J.F., Miska E.A., Thomas J.O., Allshire R.C., and Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120–124.

Barton S.C., Surani M.A., and Norris M.L. 1984. Role of paternal and maternal genomes in mouse development. *Nature* **311:** 374–376.

Bastow R., Mylne J.S., Lister C., Lippman Z., Martienssen R.A., and Dean C. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* **427:** 164–167.

Baxter J., Sauer S., Peters A., John R., Williams R., Caparros M.L., Arney K., Otte A., Jenuwein T., Merkenschlager M., and Fisher A.G. 2004. Histone hypomethylation is an indicator of epigenetic plasticity in quiescent lymphocytes. *EMBO J.* **23:** 4462–4472.

Berger S.L. 2002. Histone modifications in transcriptional regulation. *Curr. Opin. Genet. Dev.* **12:** 142–148.

Bernard P., Maure J.F., Partridge J.F., Genier S., Javerzat J.P., and Allshire R.C. 2001. Requirement of heterochromatin for cohesion at centromeres. *Science* **294:** 2539–2542.

Bernstein B.E., Kamal M., Lindblad-Toh K., Bekiranov S., Bailey D.K., Huebert D.J., McMahon S., Karlsson E.K., Kulbokas E.J., III, Gingeras T.R., et al. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120:** 169–181.

Bernstein B.E., Mikkelsen T.S., Xie X., Kamal M., Huebert D.J., Cuff J., Fry B., Meissner A., Wernig M., Plath K., et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125:** 315–326.

Bernstein E. and Allis C.D. 2005. RNA meets chromatin. *Genes Dev.* **19:** 1635–1655.

Bird A.P. 1986. CpG-rich islands and the function of DNA methylation. *Nature* **321:** 209–213.

Birky C.W., Jr. 2001. The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models. *Annu. Rev. Genet.* **35:** 125–148.

Bolland D.J., Wood A.L., Johnston C.M., Bunting S.F., Morgan G., Chakalova L., Fraser P.J., and Corcoran A.E. 2004. Antisense intergenic transcription in V(D)J recombination. *Nat. Immunol.* **5:** 630–637.

Bourc'his D. and Bestor T.H. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431:** 96–99.

Boyer L.A., Plath K., Zeitlinger J., Brambrink T., Medeiros L.A., Lee T.I., Levine S.S., Wernig M., Tajonar A., Ray M.K., et al. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* **441:** 349–353.

Bracken A.P., Dietrich N., Pasini D., Hansen K.H., and Helin K. 2006. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **20:** 1123–1136.

Braig M., Lee S., Loddenkemper C., Rudolph C., Peters A.H., Schlegelberger B., Stein H., Dorken B., Jenuwein T., and Schmitt C.A. 2005. Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* **436:** 660–665.

Brownell J.E., Zhou J., Ranalli T., Kobayashi R., Edmondson D.G., Roth S.Y., and Allis C.D. 1996. Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84:** 843–851.

Cairns B.R. 2005. Chromatin remodeling complexes: Strength in diversity, precision through specialization. *Curr. Opin. Genet. Dev.* **15:** 185–190.

Campbell K.H., McWhir J., Ritchie W.A., and Wilmut I. 1996. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380:** 64–66.

Cao R., Wang L., Wang H., Xia L., Erdjument-Bromage H., Tempst P., Jones R.S., and Zhang Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298:** 1039–1043.

Cavalli G. and Paro R. 1999. Epigenetic inheritance of active chromatin after removal of the main transactivator. *Science* **286:** 955–958.

Chakalova L., Debrand E., Mitchell J.A., Osborne C.S., and Fraser P. 2005. Replication and transcription: Shaping the landscape of the genome. *Nat. Rev. Genet.* **6:** 669–677.

Chalker D.L. and Yao M.C. 2001. Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in *Tetrahymena thermophila*. *Genes Dev.* **15:** 1287–1298.

Chan S.W., Zilberman D., Xie Z., Johansen L.K., Carrington J.C., and Jacobsen S.E.2004. RNA silencing genes control de novo DNA methylation. *Science* **303:** 1336.

Chen R.Z., Pettersson U., Beard C., Jackson-Grusby L., and Jaenisch R. 1998. DNA hypomethylation leads to elevated mutation rates. *Nature* **395:** 89–93.

Cheng D., Yadav N., King R.W., Swanson M.S., Weinstein E.J., and Bedford M.T. 2004. Small molecule regulators of protein arginine methyltransferases. *J. Biol. Chem.* **279:** 23892–23899.

Cheng J.C., Yoo C.B., Weisenberger D.J., Chuang J., Wozniak C., Liang G., Marquez V.E., Greer S., Orntoft T.F., Thykjaer T., and Jones P.A. 2004. Preferential response of cancer cells to zebularine. *Cancer Cell* **6:** 151–158.

Cho K.S., Elizondo L.I., and Boerkoel C.F. 2004. Advances in chromatin remodeling and human disease. *Curr. Opin. Genet. Dev.* **14:** 308–315.

Chuikov S., Kurash J.K., Wilson J.R., Xiao B., Justin N., Ivanov G.S., McKinney K., Tempst P., Prives C., Gamblin S.J., et al. Regulation of p53 activity through lysine methylation. *Nature* **432:** 353–360.

Clark-Adams C.D., Norris D., Osley M.A., Fassler J.S., and Winston F. 1988. Changes in histone gene dosage alter transcription in yeast. *Genes Dev.* **2:** 150–159.

Cohen F.E. and Prusiner S.B. 1998. Pathologic conformations of prion

proteins. *Annu. Rev. Biochem.* **67:** 793–819.

Cosgrove M.S., Boeke J.D., and Wolberger C. 2004. Regulated nucleosome mobility and the histone code. *Nat. Struct. Mol. Biol.* **11:** 1037–1043.

Cremer T. and Cremer C. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* **2:** 292–301.

Daujat S., Zeissler U., Waldmann T., Happel N., and Schneider R. 2005. HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. *J. Biol. Chem.* **280:** 38090–38095.

Dellino G.I., Schwartz Y.B., Farkas G., McCabe D., Elgin S.C., and Pirrotta V. 2004. Polycomb silencing blocks transcription initiation. *Mol. Cell* **13:** 887–893.

Dhalluin C., Carlson J.E., Zeng L., He C., Aggarwal A.K., and Zhou M.M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399:** 491–496.

Di Croce L. 2005. Chromatin modifying activity of leukaemia associated fusion proteins. *Hum. Mol. Genet.* **14 Spec. No. 1:** R77–R84.

Dou Y. and Gorovsky M.A. 2000. Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch. *Mol. Cell* **6:** 225–231.

Fan Y., Nikitina T., Zhao J., Fleury T.J., Bhattacharyya R., Bouhassira E.E., Stein A., Woodcock C.L., and Skoultchi A.I. 2005. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123:** 1199–1212.

Feinberg A.P. 2004. The epigenetics of cancer etiology. *Semin. Cancer Biol.* **14:** 427–432.

Feinberg A.P. and Tycko B. 2004. The history of cancer epigenetics. *Nat. Rev. Cancer* **4:** 143–153.

Feinberg A.P. and Vogelstein B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301:** 89–92.

Felsenfeld G. and Groudine M. 2003. Controlling the double helix. *Nature* **421:** 448–453.

Fischle W., Wang Y., and Allis C.D. 2003a. Binary switches and modification cassettes in histone biology and beyond. *Nature* **425:** 475–479.

———. 2003b. Histone and chromatin cross-talk. *Curr. Opin. Cell Biol.* **15:** 172–183.

Fischle W., Tseng B.S., Dormann H.L., Ueberheide B.M., Garcia B.A., Shabanowitz J., Hunt D.F., Funabiki H., and Allis C.D. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438:** 1116–1122.

Fisher A.G. and Merkenschlager M. 2002. Gene silencing, cell fate and nuclear organisation. *Curr. Opin. Genet. Dev.* **12:** 193–197.

Fodor B.D., Kubicek S., Yonezawa M., O'Sullivan R.J., Sengupta R., Perez-Burgos L., Opravil S., Mechtler K., Schotta G., and Jenuwein T. 2006. Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev.* **20:** 1557–1562.

Fraga M.F., Ballestar E., Paz M.F., Ropero S., Setien F., Ballestar M.L., Heine-Suner D., Cigudosa J.C., Urioste M., Benitez J., et al. 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci.* **102:** 10604–10609.

Francis N.J., Kingston R.E., and Woodcock C.L. 2004. Chromatin compaction by a polycomb group protein complex. *Science* **306:** 1574–1577.

Fukagawa T., Nogami M., Yoshikawa M., Ikeno M., Okazaki T., Takami Y., Nakayama T., and Oshimura M. 2004. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat. Cell Biol.* **6:** 784–791.

Garcia-Manero G. and Issa J.P. 2005. Histone deacetylase inhibitors: A review of their clinical status as antineoplastic agents. *Cancer Invest.* **23:** 635–642.

Gilbert N., Boyle S., Fiegler H., Woodfine K., Carter N.P., and Bickmore W.A. 2004. Chromatin architecture of the human genome: Gene-rich domains are enriched in open chromatin fibers. *Cell* **118:** 555–566.

Gilfillan G.D., Dahlsveen I.K., and Becker P.B. 2004. Lifting a chromosome: Dosage compensation in *Drosophila melanogaster*. *FEBS Lett.* **567:** 8–14.

Goll M.G., Kirpekar F., Maggert K.A., Yoder J.A., Hsieh C.L., Zhang X., Golic K.G., Jacobsen S.E., and Bestor T.H. 2006. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science* **311:** 395–398.

Gonzalo S. and Blasco M.A. 2005. Role of Rb family in the epigenetic definition of chromatin. *Cell Cycle* **4:** 752–755.

Gottschling D.E., Aparicio O.M., Billington B.L., and Zakian V.A. 1990. Position effect at *S. cerevisiae* telomeres: Reversible repression of Pol II transcription. *Cell* **63:** 751–762.

Grant P.A., Duggan L., Cote J., Roberts S.M., Brownell J.E., Candau R., Ohba R., Owen-Hughes T., Allis C.D., Winston F., et al. 1997. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: Characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev.* **11:** 1640–1650.

Greiner D., Bonaldi T., Eskeland R., Roemer E., and Imhof A. 2005. Identification of a specific inhibitor of the histone methyltransferase SU(VAR)3-9. *Nat. Chem. Biol.* **1:** 143–145.

Grewal S.I. and Klar A.J. 1996. Chromosomal inheritance of epigenetic states in fission yeast during mitosis and meiosis. *Cell* **86:** 95–101.

Grozinger C.M. and Schreiber S.L. 2002. Deacetylase enzymes: Biological functions and the use of small-molecule inhibitors. *Chem. Biol.* **9:** 3–16.

Guarente L. and Picard F. 2005. Calorie restriction—The SIR2 connection. *Cell* **120:** 473–482.

Hall I.M., Shankaranarayana G.D., Noma K., Ayoub N., Cohen A., and Grewal S.I. 2002. Establishment and maintenance of a heterochromatin domain. *Science* **297:** 2232–2237.

Hanahan D. and Weinberg R.A. 2000. The hallmarks of cancer. *Cell* **100:** 57–70.

Harvey A.C. and Downs J.A. 2004. What functions do linker histones provide? *Mol. Microbiol.* **53:** 771–775.

Hassan A.H., Prochasson P., Neely K.E., Galasinski S.C., Chandy M., Carrozza M.J., and Workman J.L. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111:** 369–379.

Heard E. 2005. Delving into the diversity of facultative heterochromatin: The epigenetics of the inactive X chromosome. *Curr. Opin. Genet. Dev.* **15:** 482–489.

Henikoff S. 2005. Histone modifications: Combinatorial complexity or cumulative simplicity? *Proc. Natl. Acad. Sci.* **102:** 5308–5309.

Henikoff S. and Ahmad K. 2005. Assembly of variant histones into chromatin. *Annu. Rev. Cell Dev. Biol.* **21:** 133–153.

Herr A.J., Jensen M.B., Dalmay T., and Baulcombe D.C. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308:** 118–120.

Hirota T., Lipp J.J., Toh B.H., and Peters J.M. 2005. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438:** 1176–1180.

Hochedlinger K., Blelloch R., Brennan C., Yamada Y., Kim M., Chin L., and Jaenisch R. 2004. Reprogramming of a melanoma genome by nuclear transplantation. *Genes Dev.* **18:** 1875–1885.

Holbert M.A. and Marmorstein R. 2005. Structure and activity of enzymes that remove histone modifications. *Curr. Opin. Struct. Biol.* **15:** 673–680.

Holliday R. 1994. Epigenetics: An overview. *Dev. Genet.* **15:** 453–457.

Ishii K.J. and Akira S. 2005. TLR ignores methylated RNA? *Immunity* **23:** 111–113.

Jackson J.P., Lindroth A.M., Cao X., and Jacobsen S.E. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416:** 556–560.

Jacobson R.H., Ladurner A.G., King D.S. and Tjian R. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science.* **288:** 1422–1425.

Jaenisch R. and Bird A. 2003. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nat. Genet.* (suppl.) **33:** 245–254.

Janicki S.M., Tsukamoto T., Salghetti S.E., Tansey W.P., Sachidanandam R., Prasanth K.V., Ried T., Shav-Tal Y., Bertrand E., Singer R.H., and Spector D.L. 2004. From silencing to gene expression:

Real-time analysis in single cells. *Cell* **116:** 683–698.

Jeddeloh J.A., Stokes T.L., and Richards E.J. 1999. Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nat. Genet.* **22:** 94–97.

Jenuwein T. and Allis C.D. 2001. Translating the histone code. *Science* **293:** 1074–1080.

Jones P.A. and Martienssen R. 2005. A blueprint for a Human Epigenome Project: The AACR Human Epigenome Workshop. *Cancer Res.* **65:** 11241–11246.

Kanellopoulou C., Muljo S.A., Kung A.L., Ganesan S., Drapkin R., Jenuwein T., Livingston D.M., and Rajewsky K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* **19:** 489–501.

Kayne P.S., Kim U.J., Han M., Mullen J.R., Yoshizaki F., and Grunstein M. 1988. Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell* **55:** 27–39.

Khochbin S. 2001. Histone H1 diversity: Bridging regulatory signals to linker histone function. *Gene* **271:** 1–12.

Khorasanizadeh S. 2004. The nucleosome: From genomic organization to genomic regulation. *Cell* **116:** 259–272.

Kim J., Daniel J., Espejo A., Lake A., Krishna M., Xia L., Zhang Y., and Bedford M.T. 2006. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.* **4:** 397–403.

Kimmins S. and Sassone-Corsi P. 2005. Chromatin remodeling and epigenetic features of germ cells. *Nature* **434:** 583–589.

Klar A.J. 1998. Propagating epigenetic states through meiosis: Where Mendel's gene is more than a DNA moiety. *Trends Genet.* **14:** 299–301.

———. 2004. An epigenetic hypothesis for human brain laterality, handedness, and psychosis development. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 499–506.

Knoepfler P.S., Zhang X.-Y., Cheng P.F., Gafken P.R., McMahon S.B., and Eisenman R.N. 2006. Myc influences global chromatin structure. *EMBO J.* **25:** 2723–2734.

Kornberg R.D. 1974. Chromatin structure: A repeating unit of histones and DNA. *Science* **184:** 868–871.

Lachner M., O'Sullivan R.J., and Jenuwein T. 2003. An epigenetic road map for histone lysine methylation. *J. Cell Sci.* **116:** 2117–2124.

Lachner M., Sengupta R., Schotta G., and Jenuwein T. 2004. Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 209–218.

Lachner M., O'Carroll D., Rea S., Mechtler K., and Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410:** 116–120.

Langst G. and Becker P.B. 2004. Nucleosome remodeling: One mechanism, many phenomena? *Biochim. Biophys. Acta* **1677:** 58–63.

Lee D.Y., Teyssier C., Strahl B.D., and Stallcup M.R. 2005. Role of protein methylation in regulation of transcription. *Endocr. Rev.* **26:** 147–170.

Lee T.I., Jenner R.G., Boyer L.A., Guenther M.G., Levine S.S., Kumar R.M., Chevalier B., Johnstone S.E., Cole M.F., Isono K., et al. 2006. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125:** 301–313.

Lengauer C., Kinzler K.W., and Vogelstein B. 1998. Genetic instabilities in human cancers. *Nature* **396:** 643–649.

Li E., Bestor T.H., and Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915–926.

Lippman Z., Gendrel A.V., Black M., Vaughn M.W., Dedhia N., McCombie W.R., Lavine K., Mittal V., May B., Kasschau K.D., et al. 2004. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430:** 471–476.

Litt M.D., Simpson M., Gaszner M., Allis C.D., and Felsenfeld G. 2001. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* **293:** 2453–2455.

Luger K., Mader A.W., Richmond R.K., Sargent D.F., and Richmond T.J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389:** 251–260.

Lund A.H. and van Lohuizen M. 2004. Polycomb complexes and silencing mechanisms. *Curr. Opin. Cell Biol.* **16:** 239–246.

Lyko F. 2001. DNA methylation learns to fly. *Trends Genet.* **17:** 169–172.

Maher E.R. 2005. Imprinting and assisted reproductive technology. *Hum. Mol. Genet.* **14 Spec. No. 1:** R133–R138.

Maison C., Bailly D., Peters A.H., Quivy J.P., Roche D., Taddei A., Lachner M., Jenuwein T., and Almouzni G. 2002. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* **30:** 329–334.

Marks P.A. and Jiang X. 2005. Histone deacetylase inhibitors in programmed cell death and cancer therapy. *Cell Cycle* **4:** 549–551.

Martens J.H., O'Sullivan R.J., Braunschweig U., Opravil S., Radolf M., Steinlein P., and Jenuwein T. 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **24:** 800–812.

Maurer-Stroh S., Dickens N.J., Hughes-Davies L., Kouzarides T., Eisenhaber F., and Ponting C.P. 2003. The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem. Sci.* **28:** 69–74.

Mayer W., Niveleau A., Walter J., Fundele R., and Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature* **403:** 501–502.

McClintock B. 1951. Chromosome organization and genic expression. *Cold Spring Harbor Symp. Quant. Biol.* **16:** 13–47.

McGarvey K., Fahrner J., Green E., Martens J., Jenuwein T., and Baylin S.B. 2006. Silenced tumor suppressor genes reactivated by DNA demethylation do not return to a fully euchromatic chromatin state. *Cancer Res.* **66:** 3541–3549.

McGrath J. and Solter D. 1984. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37:** 179–183.

Metzger E., Wissmann M., Yin N., Muller J.M., Schneider R., Peters A.H., Gunther T., Buettner R., and Schule R. 2005. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437:** 436–439.

Meyer B.J., McDonel P., Csankovszki G., and Ralston E. 2004. Sex and X-chromosome-wide repression in *Caenorhabditis elegans*. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 71–79.

Misteli T. 2004. Spatial positioning; a new dimension in genome function. *Cell* **119:** 153–156.

Mito Y., Henikoff J.G., and Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* **37:** 1090–1097.

Mizuguchi G., Shen X., Landry J., Wu W.H., Sen S., and Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303:** 343–348.

Mochizuki K., Fine N.A., Fujisawa T., and Gorovsky M.A. 2002. Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. *Cell* **110:** 689–699.

Mohrmann L. and Verrijzer C.P. 2005. Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta* **1681:** 59–73.

Morgan H.D., Santos F., Green K., Dean W., and Reik W. 2005. Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* **14 Spec. No. 1:** R47–R58.

Motamedi M.R., Verdel A., Colmenares S.U., Gerber S.A., Gygi S.P., and Moazed D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119:** 789–802.

Muller H.J. 1930. Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* **22:** 299–334.

Nakayama J., Rice J.C., Strahl B.D., Allis C.D., and Grewal S.I. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292:** 110–113.

Narita M., Nunez S., Heard E., Narita M., Lin A.W., Hearn S.A., Spector D.L., Hannon G.J., and Lowe S.W. 2003. Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113:** 703–716.

Narlikar G.J., Fan H.Y., and Kingston R.E. 2002. Cooperation between

complexes that regulate chromatin structure and transcription. *Cell* **108:** 475–487.

Nowak S.J. and Corces V.G. 2004. Phosphorylation of histone H3: A balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* **20:** 214–220.

Orphanides G., LeRoy G., Chang C.H., Luse D.S., and Reinberg D. 1998. FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92:** 105–116.

Pal-Bhadra M., Leibovitch B.A., Gandhi S.G., Rao M., Bhadra U., Birchler J.A., and Elgin S.C. 2004. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303:** 669–672.

Paro R. and Hogness D.S. 1991. The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl. Acad. Sci.* **88:** 263–267.

Petersen-Mahrt S. 2005. DNA deamination in immunity. *Immunol. Rev.* **203:** 80–97.

Pirrotta V. 1998. Polycombing the genome: PcG, trxG, and chromatin silencing. *Cell* **93:** 333–336.

Pontier D., Yahubyan G., Vega D., Bulski A., Saez-Vasquez J., Hakimi M.A., Lerbs-Mache S., Colot V., and Lagrange T. 2005. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. *Genes Dev.* **19:** 2030–2040.

Ratcliff F., Harrison B.D., and Baulcombe D.C. 1997. A similarity between viral defense and gene silencing in plants. *Science* **276:** 1558–1560.

Razin A. and Riggs A.D. 1980. DNA methylation and gene function. *Science* **210:** 604–610.

Rea S., Eisenhaber F., O'Carroll D., Strahl B.D., Sun Z.W., Schmid M., Opravil S., Mechtler K., Ponting C.P., Allis C.D., and Jenuwein T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406:** 593–599.

Reinberg D., Chuikov S., Farnham P., Karachentsev D., Kirmizis A., Kuzmichev A., Margueron R., Nishioka K., Preissner T.S., Sarma K., et al. 2004. Steps toward understanding the inheritance of repressive methyl-lysine marks in histones. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 171–182.

Reinhart B.J. and Bartel D.P. 2002. Small RNAs correspond to centromere heterochromatic repeats. *Science* **297:** 1831.

Rine J. 2005. Cell biology. Twists in the tale of the aging yeast. *Science* **310:** 1124–1125.

Ringrose L. and Paro R. 2004. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38:** 413–443.

Ringrose L., Ehret H., and Paro R. 2004. Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of polycomb complexes. *Mol. Cell* **16:** 641–653.

Robertson K.D. 2005. DNA methylation and human disease. *Nat. Rev. Genet.* **6:** 597–610.

Roloff T.C. and Nuber U.A. 2005. Chromatin, epigenetics and stem cells. *Eur. J. Cell Biol.* **84:** 123–135.

Roth S.Y., Denu J.M., and Allis C.D. 2001. Histone acetyltransferases. *Annu. Rev. Biochem.* **70:** 81–120.

Sanchez-Elsner T., Gou D., Kremmer E., and Sauer F. 2006. Noncoding RNAs of trithorax response elements recruit *Drosophila* Ash1 to ultrabithorax. *Science* **311:** 1118–1123.

Santos-Rosa H., Schneider R., Bannister A.J., Sherriff J., Bernstein B.E., Emre N.C., Schreiber S.L., Mellor J., and Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419:** 407–411.

Sarma K. and Reinberg D. 2005. Histone variants meet their match. *Nat. Rev. Mol. Cell Biol.* **6:** 139–149.

Scaffidi P., Gordon L., and Misteli T. 2005. The cell nucleus and aging: Tantalizing clues and hopeful promises. *PLoS. Biol.* **3:** e395.

Schalch T., Duda S., Sargent D.F., and Richmond T.J. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436:** 138–141.

Schmitt S., Prestel M., and Paro R. 2005. Intergenic transcription through a polycomb group response element counteracts silencing. *Genes Dev.* **19:** 697–708.

Schneider R., Bannister A.J., and Kouzarides T. 2002. Unsafe SETs: Histone lysine methyltransferases and cancer. *Trends Biochem. Sci.* **27:** 396–402.

Schreiber S.L. and Bernstein B.E. 2002. Signaling network model of chromatin. *Cell* **111:** 771–778.

Schwartz B.E. and Ahmad K. 2005. Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev.* **19:** 804–814.

Schwartz Y.B., Kahn T.G., and Pirrotta V. 2005. Characteristic low density and shear sensitivity of cross-linked chromatin containing polycomb complexes. *Mol. Cell Biol.* **25:** 432–439.

Sekinger E.A., Moqtaderi Z., and Struhl K. 2005. Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell* **18:** 735–748.

Seligson D.B., Horvath S., Shi T., Yu H., Tze S., Grunstein M., and Kurdistani S.K. 2005. Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* **435:** 1262–1266.

Shen X., Mizuguchi G., Hamiche A., and Wu C. 2000. A chromatin remodeling complex involved in transcription and DNA processing. *Nature* **406:** 541–544.

Shi Y., Lan F., Matson C., Mulligan P., Whetstine J.R., Cole P.A., Casero R.A., and Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119:** 941–953.

Shorter J. and Lindquist S. 2005. Prions as adaptive conduits of memory and inheritance. *Nat. Rev. Genet.* **6:** 435–450.

Sims R.J., III, Belotserkovskaya R., and Reinberg D. 2004. Elongation by RNA polymerase II: The short and long of it. *Genes Dev.* **18:** 2437–2468.

Smith C.L. and Peterson C.L. 2005. ATP-dependent chromatin remodeling. *Curr. Top. Dev. Biol.* **65:** 115–148.

Squazzo S.L., O'Geen H., Komashko V.M., Krig S.R., Jin V.X., Jang S.W., Margueron R., Reinberg D., Green R., and Farnham P.J. 2006. Suz12 binds to silenced regions on the genome in a cell-type-specific manner. *Genome Res.* **16:** 890–900.

Sterner D.E. and Berger S.L. 2000. Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* **64:** 435–459.

Strahl B.D. and Allis C.D. 2000. The language of covalent histone modifications. *Nature* **403:** 41–45.

Strahl B.D., Ohba R., Cook R.G., and Allis C.D. 1999. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proc. Natl. Acad. Sci.* **96:** 14967–14972.

Su G.H., Sohn T.A., Ryu B., and Kern S.E. 2000. A novel histone deacetylase inhibitor identified by high-throughput transcriptional screening of a compound library. *Cancer Res.* **60:** 3137–3142.

Sung S. and Amasino R.M. 2004. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature* **427:** 159–164.

Surani M.A., Barton S.C., and Norris M.L. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308:** 548–550.

Tagami H., Ray-Gallet D., Almouzni G., and Nakatani Y. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116:** 51–61.

Tamaru H. and Selker E.U. 2001. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414:** 277–283.

Tanaka E.M. 2003. Regeneration: If they can do it, why can't we? *Cell* **113:** 559–562.

Thomas J.O. 1999. Histone H1: Location and role. *Curr. Opin. Cell Biol.* **11:** 312–317.

Tsukada Y., Fang J., Erdjument-Bromage H., Warren M.E., Borchers C.H., Tempst P., and Zhang Y. 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439:** 811–816.

Tsukiyama T., Daniel C., Tamkun J., and Wu C. 1995. ISWI, a member of the SWI2/SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. *Cell* **83:** 1021–1026.

Turner B.M. 2000. Histone acetylation and an epigenetic code. *BioEssays* **22:** 836–845.

Valk-Lingbeek M.E., Bruggeman S.W., and van Lohuizen M. 2004. Stem cells and cancer; the polycomb connection. *Cell* **118:** 409–418.

van Attikum H. and Gasser S.M. 2005. The histone code at DNA breaks: A guide to repair? *Nat. Rev. Mol. Cell Biol.* **6:** 757–765.

van der Heijden G.W., Dieker J.W., Derijck A.A., Muller S., Berden J.H., Braat D.D., van der Vlag J., and de Boer P. 2005. Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122:** 1008–1022.

Vaquero A., Loyola A., and Reinberg D. 2003. The constantly changing face of chromatin. *Sci. Aging Knowledge Environ.* **2003:** RE4.

Varga-Weisz P.D., Wilm M., Bonte E., Dumas K., Mann M., and Becker P.B. 1997. Chromatin-remodeling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388:** 598–602.

Verdel A., Jia S., Gerber S., Sugiyama T., Gygi S., Grewal S.I., and Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303:** 672–676.

Vidanes G.M., Bonilla C.Y., and Toczyski D.P. 2005. Complicated tails: Histone modifications and the DNA damage response. *Cell* **121:** 973–976.

Vignali M., Hassan A.H., Neely K.E., and Workman J.L.. 2000. ATP-dependent chromatin-remodeling complexes. *Mol. Cell. Biol.* **20:** 1899–1910.

Vire E., Brenner C., Deplus R., Blanchon L., Fraga M., Didelot C., Morey L., Van E.A., Bernard D., Vanderwinden J.M., et al. 2005. The Polycomb group protein *EZH2* directly controls DNA methylation. *Nature* **439:** 861–874.

Volpe T.A., Kidner C., Hall I.M., Teng G., Grewal S.I., and Martienssen R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297:** 1833–1837.

Waddington C.H. 1957. *The strategy of the genes.* MacMillan, New York.

Wade P.A., Gegonne A., Jones P.L., Ballestar E., Aubry F., and Wolffe A.P. 1999. Mi-2 complex couples DNA methylation to chromatin remodeling and histone deacetylation. *Nat. Genet.* **23:** 62–66.

Walsh C.P., Chaillet J.R., and Bestor T.H. 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20:** 116–117.

Watanabe Y., Yokobayashi S., Yamamoto M., and Nurse P. 2001. Pre-meiotic S phase is linked to reductional chromosome segregation and recombination. *Nature* **409:** 359–363.

Watson J.D. 2003. Celebrating the genetic jubilee: A conversation with James D. Watson. Interviewed by John Rennie. *Sci. Am.* **288:** 66–69.

Wei Y., Yu L., Bowen J., Gorovsky M.A., and Allis C.D. 1999. Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell* **97:** 99–109.

Whetstine J.R., Nottke A., Lan F., Huarte M., Smolikov S., Chen Z., Spooner E., Li E., Zhang G., Colaiacovo M., and Shi Y. 2006. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125:** 467–481.

Wolffe A.P. and Matzke M.A. 1999. Epigenetics: Regulation through repression. *Science* **286:** 481–486.

Wysocka J., Swigut T., Milne T.A., Dou Y., Zhang X., Burlingame A.L., Roeder R.G., Brivanlou A.H., and Allis C.D. 2005. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121:** 859–872.

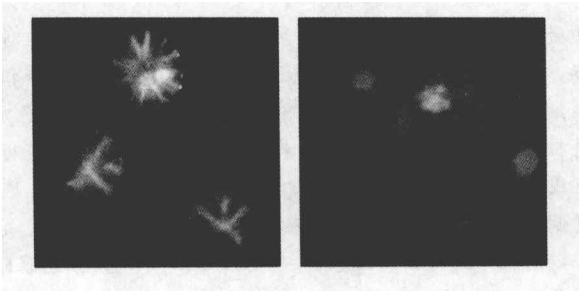
Yan Q., Huang J., Fan T., Zhu H., and Muegge K. 2003. Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. *EMBO J.* **22:** 5154–5162.

Yu B., Yang Z., Li J., Minakhina S., Yang M., Padgett R.W., Steward R., and Chen X. 2005. Methylation as a crucial step in plant microRNA biogenesis. *Science* **307:** 932–935.

Zhang Y. and Reinberg D. 2001. Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15:** 2343–2360.

Zhang Y., LeRoy G., Seelig H.P., Lane W.S., and Reinberg D. 1998. The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities. *Cell* **95:** 279–289.

# 第 4 章



# 酿酒酵母的表观遗传学

Michael Grunstein[1] and Susan M. Gasser[2]

[1]*University of California, Los Angeles, California 90095-1570*

[2]*Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland*

## 本章目录

1. 研究酵母用的遗传学和分子生物学方法
2. 酵母的生命周期
3. 酵母异染色质存在于沉默的 *HM* 接合型位点和端粒处
4. 异染色质可以通过蔓延到整个沉默区域的抑制性结构得以区分
5. 异染色质组装的不同步骤
   - 5.1 *HM* 异染色质
   - 5.2 端粒异染色质
6. Sir2 催化的组蛋白去乙酰化为 SIR 复合体蔓延提供结合位点
7. Sir2 催化组蛋白 H4 第 16 位赖氨酸的去乙酰化
8. 常染色质上的组蛋白乙酰化阻止 SIR 复合体的蔓延
9. 端粒成环
10. 端粒成环介导的天然亚端粒元件处的抑制不连续性
11. 端粒间的反式相互作用和异染色质在核周的附着
12. 表观遗传状态的继承
13. 衰老与 Sir2：因 rDNA 重复序列的不稳定性而相联系
14. 总结

参考文献

## 概要

真核生物细胞核中带有活化基因的一部分染色质被称为常染色质。这种染色质在有丝分裂时被压缩以使染色体分离，而在细胞周期的间期解压缩以促进转录的发生。但是，根据细胞形态观察，有些染色体区域在细胞周期的间隙仍然是压缩的，这种组成型压缩染色质称为异染色质。随着新技术的发展，分子特征取代了细胞学特征被用来对基因组上这段区域进行定义，这种位于着丝粒和端粒中的组成型压缩染色质被发现包含有简单序列数以千计的重复。这种异染色质倾向于在细胞周期的S期晚期复制并且聚集在核周边或者核仁附近区域。重要的是，它特征性的核酸酶抗性染色质结构可以蔓延并以随机方式抑制周围基因。例如，果蝇中决定红色眼睛的“白眼”基因，由于一种位置效应花斑现象，受到表观遗传抑制而产生红白镶嵌的眼睛。从机制上，位置效应花斑反映HP1蛋白对组蛋白H3K9甲基化的识别以及这一标记沿着染色体臂的蔓延。在酿酒酵母（即酿酒酵母）中，进化出了一种不同的异染色质形成机制，但其结果是相似的。

酿酒酵母是一种常用于啤酒酿造及面包焙烤的微生物。但是，与细菌不同的是，它是一种真核生物。酿酒酵母的染色体，和复杂真核生物的染色体一样，包裹着组蛋白，被封闭在核内，并且在细胞周期的S期，从多个起点开始复制。然而酵母基因组非常小，由分布于16条染色体的$1.4\times10^7$对碱基组成，其中有些染色体甚至比某些噬菌体的基因组还小。在酵母基因组中，大约有6000个基因，沿着染色体臂紧紧相连，相邻基因间的距离通常小于2 kb。酵母的大多数基因都处于开放的染色质状态，也就是说它们或者处于非常活跃的转录中，或者可以很快被诱导。另外酿酒酵母中只有非常有限数量的简单DNA重复序列，上述因素使得酿酒酵母中的异染色质无法被细胞学技术所检测。

然而，分子生物学的技术已经证实，酿酒酵母所有16条染色体上临近端粒处以及Ⅲ号染色体上两个沉默接合型位点处有着明显的类似异染色质的区域。对两个沉默接合型位点的转录抑制为维持具交配能力的单倍体状态所必需。亚端粒区域和沉默接合型位点都能以位置依赖的表观遗传调控来抑制整合进去的报告基因。它们在S期晚期开始复制并且存在于核周边。因此，虽然这些位点不能形成细胞间期可见的压缩状态，但是它们拥有许多异染色质的特征。的确，对于研究异染色质的科学家来说，酵母有很多优点，它的基因组小，有可利用的研究微生物的遗传和生化方法，并且具备高级真核生物染色体的重要特征。

## 1. 研究酵母用的遗传学和分子生物学方法

酵母提供了一种灵活且快速的遗传体系，可用于对细胞内事件的研究。酵母的一代约 90min，包含数百万个细胞的菌落在两天之内就可以长成。除此之外，酵母能以单倍体和二倍体方式繁殖，极大地促进了遗传学分析。类似于细菌，单倍体酵母细胞可以被突变为营养缺陷型。隐性致死突变能以条件致死等位基因的单倍体（如温度敏感型突变体）或者杂合二倍体（同时带有野生型和突变的等位基因）的方式保存。酵母中高效的同源重组系统使得我们可以对任一选定的染色体序列进行改变。而且，染色体片段可以被重组到质粒上，只需添加含有着丝粒和 DNA 复制起始点的简单序列就可以在酵母细胞分裂时被稳定地保留。甚至可以被重组到线性质粒或微型染色体，只要在末端添加端粒重复序列它们也可以在酵母中稳定存在。

以果蝇“白眼基因”作为报告基因的位置效应花斑在确认表观遗传基因调控和影响这一独特基因抑制形式的基因中显得尤为重要（详见第 5 章）。借助于 *Ura3* 和 *Ade2* 报告基因，在酵母端粒附近发现了一个相似现象，即端粒位置效应（telomere position effect，TPE）（图 4-1）。当 5-氟乳清酸（5-FOA）存在时，Ura3 蛋白可以将 5-FOA 改

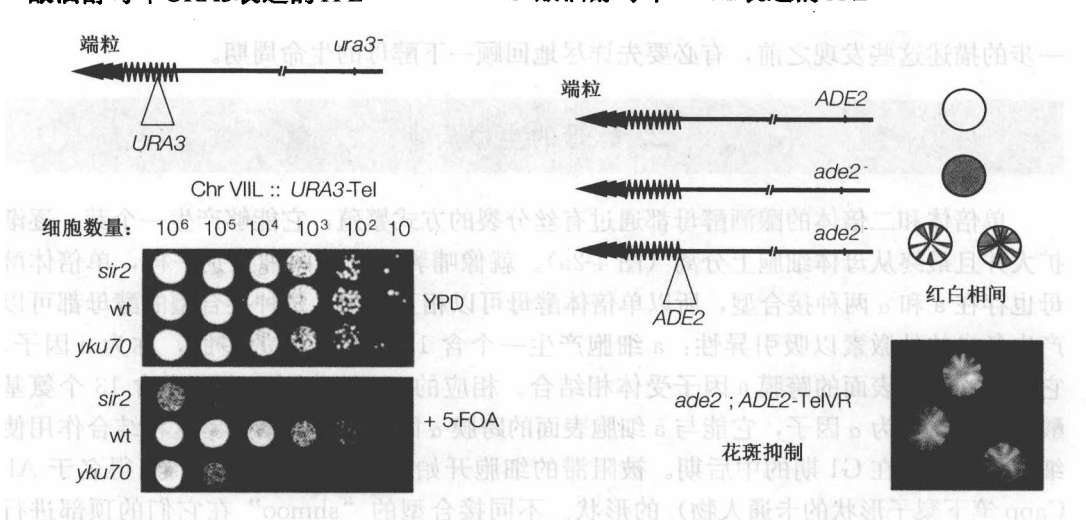



图 4-1　酵母中的基因沉默和端粒位置效应

(a) 插入到酵母染色体 VII 左臂附近的简单 TG 重复处的 *Ura3* 基因被端粒异染色质所沉默。培养于一般的丰富培养基（YPD）时，*Ura3* 基因被抑制的野生型细胞和 *Ura3* 基因表达的突变体（失去了端粒异染色质）之间没有生长上的差异。但在含有 5-FOA 的培养基中只有 *Ura3* 被抑制的细胞才能生长，而表达 *Ura3* 的细胞（*Sir2* 和 *yku70* 的突变体）不能生长。因为 *Ura3* 基因的产物可以把 5-FOA 转变成有毒的中间产物 5-氟尿嘧啶。通过浓度梯度试验可以检测到低至每 $10^6$ 细胞中产生的一个沉默细胞。(b) 含有野生型 *Ade2* 基因的细胞产生白色的菌落，而含有突变 *Ade2* 基因的细胞却显红色，因为细胞累积了腺嘌呤生物合成途径中的一种红色中间产物。当 *Ade2* 基因被插入到染色体 V 右臂邻近端粒处，它通过一种表观遗传的方式发生沉默。在遗传上相同的细胞中，*Ade2* 的沉默或活跃状态均可被继承，导致在同一菌落中产生白色或红色的扇形区域（很像位置效应花斑）。

变为 5-氟尿嘧啶（5-FU）。5-FU 是一种导致细胞死亡的 DNA 合成抑制剂。但是，当 *Ura3* 整合到异染色质区域时，*Ura3* 基因在部分细胞中被抑制，只有那些使 *Ura3* 沉默的细胞才可以在 5-FOA 的存在下生长。因此，通过测定酵母在不同浓度梯度的 5-FOA 中的生长效率，，可以在很大范围（如 10～$10^6$倍）内量化报告基因的抑制作用。此外，通过检测对 5-FOA 敏感度的增加，破坏端粒位置效应的突变可以很容易地被识别。

类似的，当把 *Ade2* 作为靶子整合到异染色质区域时，此基因会被抑制并在细胞中积累腺嘌呤生物合成的前体，呈现微红色。重要的是，*Ade2* 抑制的表观遗传特征可在遗传上相同的单一菌落中被分辨：该基因在有些细胞中以“开”的状态存在，而在另一些细胞中以“关”的状态存在，这就能在白色克隆背景中产生红色区域，反之亦然（图 4-1b）。与应用 *Ura3* 的检测方法不同的是，对 *Ade2* 基因不表达的细胞没有选择压力，因此，用插入亚端粒异染色质中的 *Ade2* 报告基因的表型不仅能体现其状态间的转换率，还能体现其表观遗传状态的可遗传性。*Ade2* 颜色反应有力地证明了抑制和非抑制状态都具有相对稳定性。

结合以上遗传学手段，生物化学技术被用在大规模培养的同步或非同步生长的蛋白酶缺陷菌株。近年来，实验手段极大发展，复杂的基因芯片和蛋白质网络技术，很快因适用于酵母小规模的基因组而被广泛采用。这些方法使得全基因组水平的转录，转录因子结合，组蛋白修饰，蛋白质-蛋白质相互作用的分析成为可能。这一系列精细的方法使得科学家能探索酿酒酵母中异染色质建立及其生理作用的调节机制。但是，在我们进一步的描述这些发现之前，有必要先详尽地回顾一下酵母的生命周期。

## 2. 酵母的生命周期

单倍体和二倍体的酿酒酵母都通过有丝分裂的方式繁殖，它能够产生一个芽，逐渐扩大并且最终从母体细胞上分离（图 4-2a）。就像哺乳动物的两种性别一样，单倍体酵母也存在 a 和 α 两种接合型，所以单倍体酵母可以相互交配。每种接合型的酵母都可以产生各自的外激素以吸引异性：a 细胞产生一个含 12 个氨基酸的多肽，称为 a 因子，它能与 α 细胞表面的跨膜 a 因子受体相结合。相应的，α 细胞则产生一个含 13 个氨基酸的多肽，称为 α 因子，它能与 a 细胞表面的跨膜 α 因子受体相结合。这种结合作用使细胞被阻滞在在 G1 期的中后期。被阻滞的细胞开始形成类似“shmoo”（得名于 ALCapp 笔下梨子形状的卡通人物）的形状。不同接合型的“shmoo”在它们的顶部进行融合，产生 a/α 二倍体。在二倍体中交配应答是被抑制的，它们和单倍体一样，进行无性繁殖（如通过有丝分裂）。另一方面，饥饿会导致减数分裂的进行，从而形成包含 4 个孢子的子囊，每种接合型两个。一旦有足够的营养，单倍体孢子会成熟为拥有交配能力的细胞，重新开始生命周期。

虽然在实验研究中单倍体酵母被认为具有确定的接合型。但在自然界中，酵母每经历一个生命周期就会改变一次接合型（图 4-3a）。接合型转变是由一种核酸内切酶（HO）所诱发，HO 能在 MAT 位点处产生位点特异性双链断裂。随后基因转换事件将组成型沉默的供体位点 HMLα 或 HMRa 置入活化的 *MAT* 位点，实现接合型的互换。

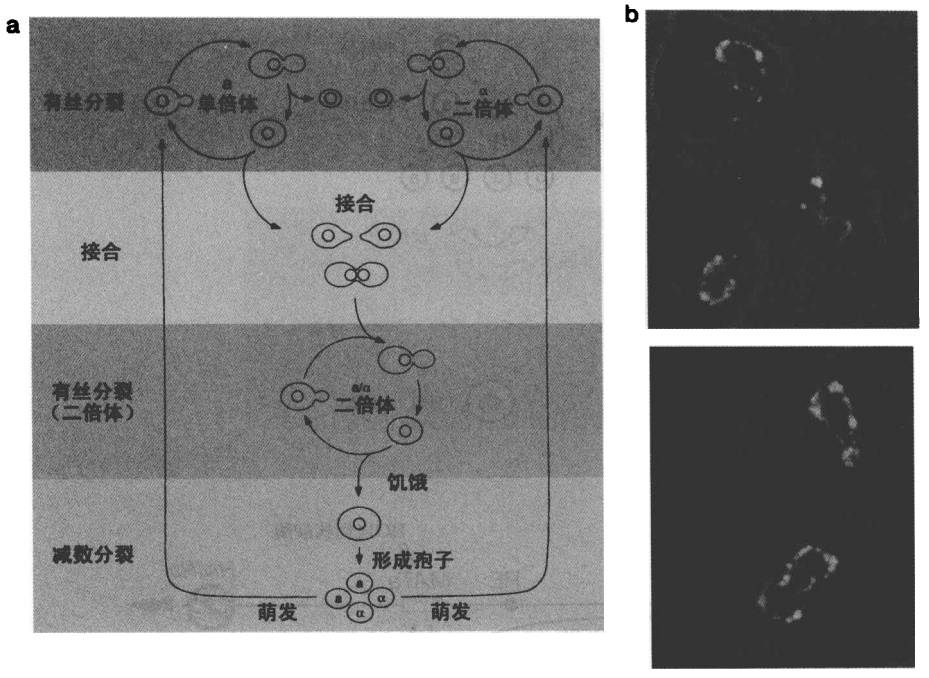



图 4-2　芽殖酵母的生命周期

(a) 单倍体或二倍体酵母均可发生有丝分裂。二倍体酵母在饥饿状态下可以诱导形成孢子，当不同接合型的单倍体临近时，接合能自发进行。接合依赖于外激素的分泌，其能使相反接合型的细胞停滞在 $G_1$ 期，在接触外激素足够的时间之后，接合的通路被诱导。在二倍体状态时接合被抑制。(b) 外激素刺激的应合，单倍体细胞扭向接合型相反的细胞。这被称为“shmoos”。核膜以绿色荧光标记。

此种菌株称作同宗配合（homothallic）的。意为一个无性繁殖的 *MAT*a 细胞可以迅速产生 *MAT*α 后代，反之亦然。因为在实验室中大家都希望细胞能有一个稳定的接合型，所以通常实验室中的菌株被构建成含有一个 HO 核酸内切酶基因突变，而失去切开 *MAT* 位点的能力。HO 内切酶的缺失阻止了不同接合型之间的转换，产生了一个异宗配合（heterothallic）的株系。这种菌株含有沉默的 *HM* 位点，和一个活化的 *MAT* 位点，其接合型信息为稳定的 a 或 α。两个沉默的“接合型”位点，分别对应着两种性别，它们通过表观遗传的方式来维持组成型的沉默，并且已经成为研究异染色质的经典模型。

### 3. 酵母异染色质存在于沉默的 *HM* 接合型位点和端粒处

三个接合型位点 *HML*α、*MAT* 和 *HMR*a 位于三号染色体上，它们含有决定酵母接合型 a 或 α 的信息。*HMR*a（距离左边端粒约 11 kb）*HMR*α（距离右边端粒约 23 kb；图 4-3b，c）定位于被称为 E 沉默子和 I 沉默子的短 DNA 元件之间。仅仅当任意一个沉默盒被复制并整合到有活性的 *MAT* 位点的时，它才具有转录能力。*HML*α 转

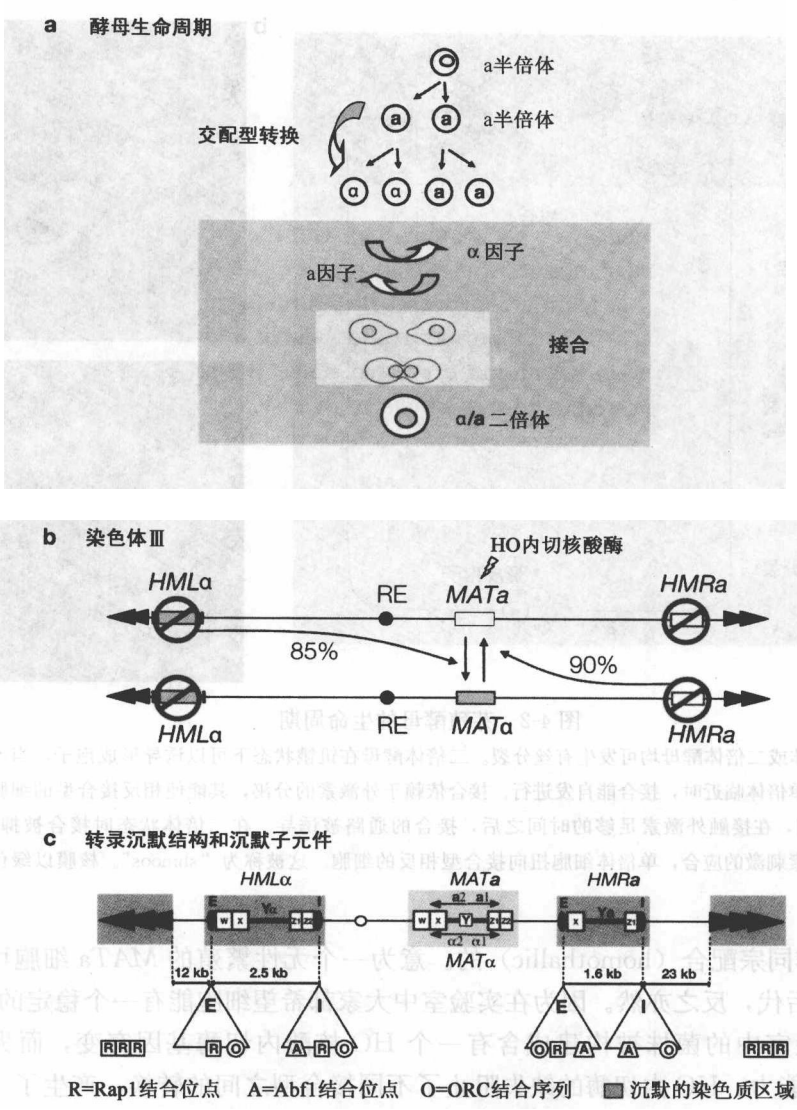



图 4-3 酵母中的接合型转换

(a) 同宗接合的酵母株经过一次分裂周期之后能转换接合型。这种转换在 DNA 复制之前发生，所以母细胞和子细胞都拥有新的接合型。(b) 染色体Ⅲ上沉默和表达的 *MAT* 位点。在一个细胞周期中有活性的 *MAT* 位点大约可以发生一次转换，该转换因 HO 内切核酸酶引起的双链断裂所致。图示的百分率表示用相反的接合型取代 *MAT* 位点的基因转换事件发生的频率。转换的方向由位于染色体Ⅲ左臂上的重组增强子所引导。(c) 沉默接合型位点 *HMR* 和 *HML* 的抑制由沉默基因两侧的沉默子 DNA 元件所介导。它们叫做 E (essential，必需的) 和 I (important，重要的) (Brand et al. 1997)，并且为 Rap1 (R)、Abf1 (A) 和 ORC (O) 提供结合位点。应用各种冗余的结合位点的组合可以构建人工的沉默子，尽管其效率不如天然的沉默子。*HMLα* 和 *HMR* 距离染色体Ⅲ的端粒分别为 12 kb 和 23 kb。染色体Ⅲ的端粒异染色质结构域的沉默是独立于 *HM* 位点的，多个 Rap1 结合位点 (R) 起始这一过程。

入 *MAT* 位点产生 α 接合型的细胞，*MHRa* 转入 *MAT* 位点则产生 a 接合型的细胞。这表明在 *HM* 处的基因和启动子是完整无缺的，虽然它们位于 *HMR* 和 *HML* 时仍被稳

定抑制。这是维持交配能力所必需，因为在一个细胞中同时表达 a 和 α 的转录产物会造成无法交配的不育状态。用不育表型来筛选破坏 HM 位点沉默状态的基因是非常有用的遗传手段。通过这种方法，沉默信息调节蛋白 SIR1、SIR2、SIR3 和 SIR4 被鉴定出来，它们被认为对于 *HM* 位点的抑制状态是必需的（Rusche et al. 2003）。*sir2*、*sir3* 和 *sir4* 的突变会导致沉默的完全丢失，然而在 *sir1* 的突变体中，仅仅一部分 *MAT*a 细胞因失去 *HM* 位点的抑制而不能交配。利用 *sir1* 缺陷细胞出现的这种部分表型，证明了在遗传上相同的细胞中可以出现两种可稳定遗传的状态（可交配和不可交配）（Pillus and Rine 1989）。这清楚的证明了接合型的抑制具有表观遗传学调控的特征。另外，别的研究证明了组蛋白 H3 和 H4 的 N 端、阻遏蛋白激活蛋白（RAP1）和起始位点识别复合体（ORC）都是异染色质的结构成分（Rusche et al. 2003）。

异染色质也存在于邻近酵母端粒末端 DNA（$C_{1\text{-}3}A/TG_{1\text{-}3}$）的位置上。如前所述，当报告基因 *Ura3* 或者 *Ade2* 被整合到这些端粒重复的邻近区域，它们出现抑制的表观遗传花斑效应（Gottschling et al. 1990）。这种端粒位置效应和 HM 一样需要 *Rap1*、*Sir2*、*Sir3*、*Sir4* 和组蛋白 N 端（Kayne et al. 1988；Aparicio et al. 1991）的参与。这从遗传学上证明除了 *Sir1*，相似的机制导致了 *HM* “接合型” 位点和端粒附近的基因沉默。此外，端粒报告基因以可检测的比率在沉默和表达之间转换，这种基因的抑制形式和果蝇的位置效应花斑十分相似。

在酵母中，这四种介导基因表达抑制的 sir 蛋白彼此之间没有广泛的同源性。并且，Sir1、Sir3、Sir4 蛋白仅在酿酒酵母以及密切相关的其他酿酒酵母中保守存在。另一方面，Sir2 是一个 NAD 依赖的组蛋白去乙酰化酶庞大家族的重要成员，它从细菌到人都是保守的（图 4-4）。Sir2 类似的组蛋白去乙酰化酶在转录抑制中的作用在别的生物体中也被发现，例如裂殖酵母和果蝇，它们没有其他的 Sir 蛋白。裂殖酵母的 Sir2 活性对于端粒附近的转录沉默是必需的，在果蝇中 Sir2 影响位置效应花斑的稳定性（chopra 和 Mishra 2005）。NAD 水解和 Sir2 的去乙酰化作用产生了 O-乙酰基 ADP 核糖，一种可能有独立作用的中间产物（Tanner et al. 2000；另见 13 章）。值得一提的是，Sir2 家族的酶修饰许多组蛋白之外的底物，其中一大分支其实是细胞质内的酶。为了应答外界压力以及 DNA 损伤，哺乳动物的 Sir2 能使转录因子 ROXO、p53 去乙酰化，激活其作用，阐明了 Sir2 功能的多样性。在酿酒酵母中，Sir2 除了基因沉默之外，还有一个重要作用，即抑制核仁中高度重复的 rDNA 基因的重组（Gottlieb and Esposito 1989）。

### 4. 异染色质可以通过蔓延到整个沉默区域的抑制性结构得以区分

常染色质基因的抑制归因于抑制蛋白或复合体识别启动子处的特殊序列，阻止了转录机器的运动及结合。异染色质的抑制通过另一种不同的非启动子专一的机制：抑制作用起始于特定位点，然后连续地蔓延至整个区域，使得该区域内所有的启动子都被沉默。这被染色质免疫共沉淀所明示，此方法显示了 Sir2、Sir3、Sir4 蛋白与整个亚端粒区沉默染色质的相互作用（Hecht et al. 1996；Strahl-Bolsinger et al. 1997）。来源于其

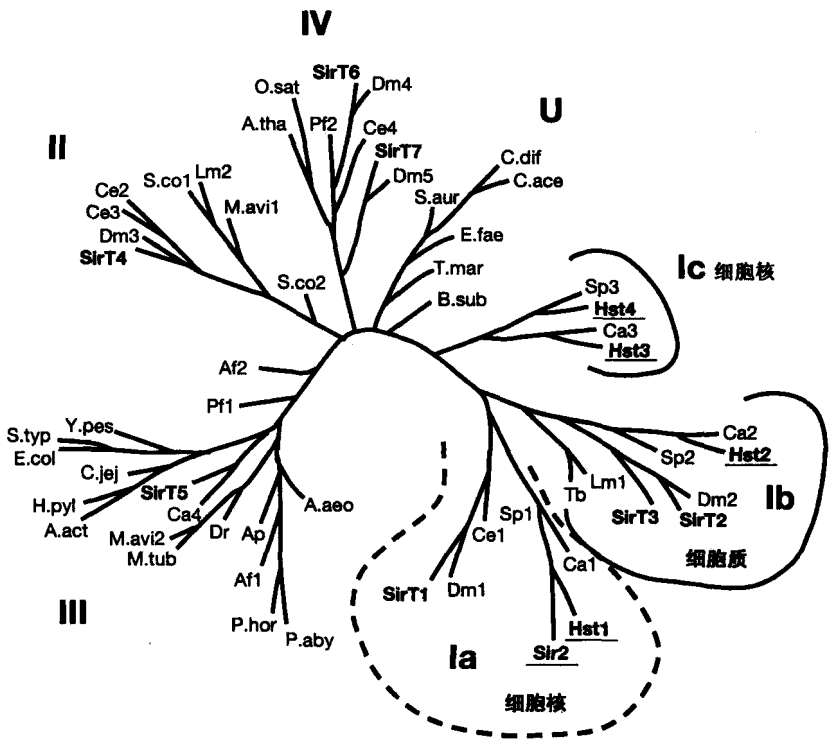



图 4-4 Sir2 去乙酰化酶家族

Sir2 是一个庞大的 NAD 依赖的组蛋白去乙酰化酶家族的重要成员。Sir2 家族蛋白通常是保守的，他们在从细菌到人类的很多有机体中被发现，并且包含了核内和胞质内的进化树上的分支。用 CLUSTAL W®和 TREEVIEW®比较 cDNA 文库和特殊序列文库中同源体的核心结构域序列，生成了 Sir2 同源体系统发生的无根树。含有六个子类和不相关组（U）。哺乳动物中的同源体 SirT1-7 用黑体标出，酿酒酵母中的蛋白用下划线标出，别的物种的蛋白用物种名字指明（经 Frye 2000 [©Elsevier] 许可，有修改）。

他一些方法的证据表明这诱导产生了抑制态的、难以被接触的染色质结构。例如，在表达细菌 *dam* 甲基化酶的酵母细胞中，尽管该酶可修饰沉默区域以外的序列，但是沉默染色质的 DNA 并没有被有效的甲基化。这说明异染色质可以限制诸如 *dam* 转甲基酶等大分子的进入（Gottschling 1992）。类似的，分离出的核中约 3kb 的 *HMR* 位点对特定的核酸内切酶有抵制作用（Loo and Rine 1994），核小体密集存在于两个沉默元件，造成在 *HM* 沉默位点出现核酸酶抗性区域，而在 *HM* 活性位点则没有（Weiss and Simpson 1998）。因此，酵母异染色质明显具有独特的染色质结构。

酵母或多细胞动物异染色质高度压缩的程度，以及压缩对转录因子带来的空间位阻的程度，还不是很确定。令人惊讶的是，由 Sir 蛋白与组蛋白形成的抑制复合物似乎是动态的，因为即使细胞周期被阻滞于异染色质装配根本不发生的时期，Sir 蛋白也可以被掺入 HM 沉默染色质（Cheng and Gartenberg 2000）。这也许可以解释为什么结合了 Sir 蛋白的异染色质在处于抑制状态的时候依旧可以作为某些转录因子的结合位点（Sekinger and Gross 1999）。虽然此类研究认为异染色质并不能阻止所有非组蛋白的接

近，但是在异染色质区域并没有明显的转录发生，也不能通过实验来检测到工作中的 RNA 聚合酶。Chen 和 Widom 的实验（2005）认为异染色质可以阻止包含了 RNA 聚合酶 II 及结合启动子的转录因子 TFIIB 和 TFIIE 的蛋白复合体的招募。因此，虽然沉默的酵母染色质是动态变化的，允许 SIR 因子和某些转录因子的交换，但是它选择性地阻止基本转录机器的结合从而阻止 mRNA 的产生（参见第 10 章）。

## 5. 异染色质组装的不同步骤

所有物种的异染色质组装都包括了一系列的分子事件，其中一些是在芽殖酵母中被发现的。位点特异的异染色质成核现象是研究地最清楚的事件之一，其过程需要序列专一性的 DNA 结合因子。随后，异染色质从起始位点开始蔓延。这种蔓延被限制在特定的边界处，其机制在酵母中已经被研究清楚。另外，酵母可以用来研究核内区室在异染色质介导的基因表达抑制中所起的作用。端粒处的异染色质的组装和在 *HM* 位点处有些不同，但是它们都反映出一个相似的原则，那就是通用抑制蛋白的蔓延以专一性 DNA 结合因子的存在为核心。下面我们来仔细描述这些机制（图 4-6）。

### 5.1 *HM* 异染色质

沉默的接合型位点 *HML* 和 *HMR* 被两个叫做沉默子 E（essential 必需的）和 I（important 重要的）的短 DNA 元件括起来（图 4-3b，c）。沉默子元件至少为两个叫做 Rap1 和 Abf1 的多功能核因子，以及复制起始点识别复合体提供了结合位点（Brand et al. 1987）。虽然缺失 *HMR*-E（在所有三个位点都存在）比缺失 *HMR*-I（存在于两个位点）对沉默表形产生更大的影响，但是每个沉默子都能作为成核位点来招募 Sir 沉默复合体，进而促使后续的 Sir 蛋白沿着他们之间的核小体蔓延。结合于两个不同位点的 Rap1 分子在体外实验中的相互接触，指出了结合在 E 和 I 处的因子可以发生直接的相互作用从而形成环形的抑制区域，这解释了为什么 E 和 I 在抑制起始中有协同作用（Hofmann et al. 1989）。

沉默子元件的冗余是异染色质介导的基因抑制的一个特点，这对于同一沉默子也是如此。Rap1 和 Abf1 是基本转录因子，ORC 则结合在复制起始点处以形成复制前复合体，它们以冗余的方式进行基因抑制。这已经被缺失实验所证明：只要有任意两个因子的结合位点就可以引起基因沉默（Brand et al. 1987）。尽管 Rap1、Abf1 和 ORC 几乎没有结构上的相似性，但是通过分析他们所招募的蛋白可以解释它们功能上的冗余性。例如，Rap1 可以在 *HM* 沉默子和端粒处招募 Sir4，Abf1 可以和 Sir3 相互作用，ORC 和 Sir1（一种对 *HM* 的抑制有专一性的 Sir 因子）蛋白有很高的亲和力（Rusche et al. 2003）。Sir1 可以直接和 Sir4 的 N 端相互作用，在 ORC 和 SIR2-3-4 复合体之间提供了桥梁。因此，不同的沉默子结合因子都可以招募 Sir4 并依次招募 SIR2-3-4 复合体，后者为所有的抑制所必须。Rap1、Abf1 和 ORC 在沉默子处功能上的冗余，可以归因于它们每个都可以直接结合 SIR 复合体的不同组分，Ku 异源二聚体（在端粒处，见下）也是这样。

值得注意的是 Sir1 主要在异染色质的建立中发挥作用，而不是其维持。一旦 Sir1 帮助建立了沉默，在这种抑制状态的维持中它将不再被需要（Pillus and Rine 1989）。Sir1 的重要作用被以下实验证明，将 Sir1 和 Gal4 的 DNA 结合结构域形成一个融合蛋白（GBD-Sir1），并用 Gal4 结合位点取代 HMR-E 沉默子，在这种情况下 GBD-Sir1 可以有效地作为沉默核心，而不再需要沉默子及其结合因子（Chien et al. 1993）。尽管如此，Sir1 靶向的抑制仍需要所有其余的 Sir 蛋白以及完整的组蛋白尾部。这说明沉默子结合因子一个基本的功能是吸引 Sir1，进而招募其余的 Sir 蛋白与邻近的核小体相互作用来成为沉默核心。有一个证据支持这个观点，与其他的 sir 蛋白不同，Sir1 并不随着 Sir 复合体蔓延到沉默子之外（图 4-5）（Rusche et al. 2002）。

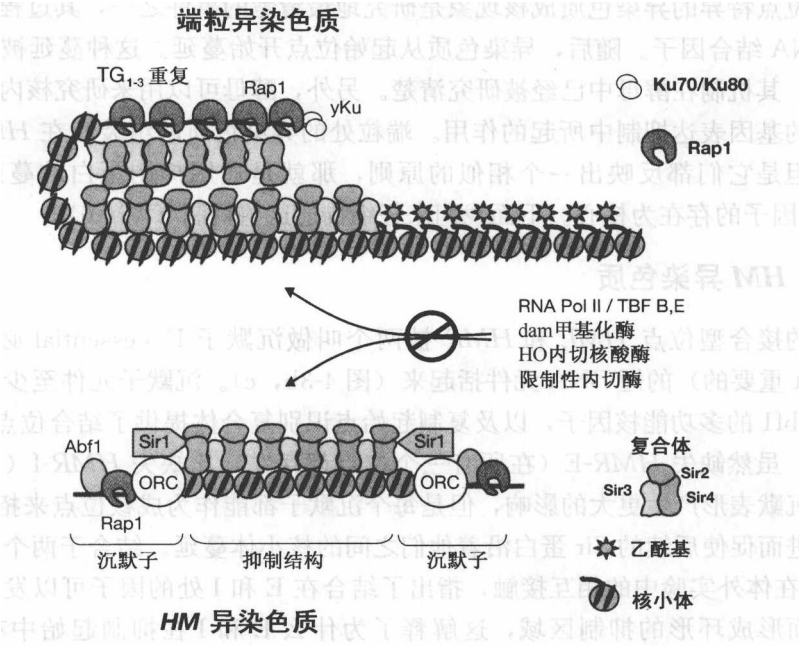



图 4-5 酵母端粒处和 HM 位点处异染色质的模型

端粒和 *HM* 位点促使 SIR 复合体蔓延的沉默子共同使用的因子包括 Rap1、Sir2、Sir3 和 Sir4，不同的是，端粒还依赖 yKu，而 *HM* 位点还使用了 ORC、Abf1 和 Sir1。端粒异染色质被认为可以折回形成一个帽子结构，以保护其免遭降解，端粒的压缩和折叠导致基因沉默。对于 *HM* 异染色质，在沉默子元件之间的抑制区域由紧挨着的核小体组成，形成一个压缩态结构。端粒和 *HM* 沉默区域都不能被转录机器和降解酶类所接近。

## 5.2 端粒异染色质

在端粒处，一个以 RNA 为基础的端粒酶维持一个约 300～350 bp 简单但不规则的富含 TG 的重复序列的长度，它为 Rap1 提供了 16～20 个一致的结合位点。一系列 Rap1 结合位点在染色体末端组成了一个没有核小体的帽子，它在端粒长度维持中扮演着重要角色（Marcand et al. 1997）。沿着端粒重复序列，Rap1 通过一个核心 DNA 结合结构域和它的结合位点结合，并且通过其羧基端结构域和 Sir4 结合，即便不存在其他 Sir 蛋白或 H4 的 N 端。因为破坏 Sir4 可以阻止其他蛋白结合到端粒染色质上（Luo

et al. 2002)，它看来似乎是成核事件与随即发生的沉默性染色质结构形成之间的关键联系。

DNA 末端结合复合体 yKu70/yKu80 帮助 Rap1 把 Sir4 招募到染色体末端。事实上，yKu 的缺失强烈地去抑制（derepress）端粒的沉默，靶定 GBD-yKu 融合蛋白可以有效地抑制一个失去沉默子的报告基因。敲除 Rap1 结合因子 Rif1 可以绕开对 yKu 的需要，Rif1 与 Rap1 羧基端结构域竞争和 Sir4 的结合（图 4-6）（Mishra and Shore 1999)。含 30 个以上的 Rap1 结合位点的 600 bp 端粒重复序列不足以在染色体内部的区域诱发沉默，但含约 45 个 Rap1 结合位点的 900 bp 序列却可以，这证明 yKu 和 Rap1

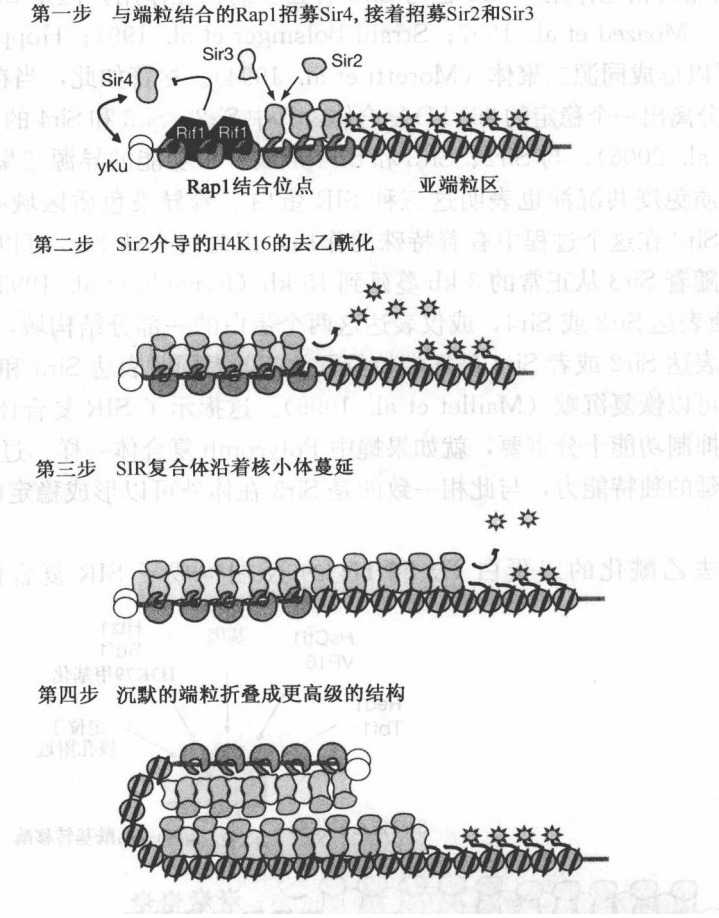



图 4-6　端粒异染色质装配的步骤

第一步：在端粒，Rap1 和 yKu 可以招募 Sir4。即便没有 Sir2 和 Sir3，Sir4 也可以在没有其他 Sir 蛋白时被招募。Sir4 的结合与 Rif1 和 Rif2 相拮抗（Mishra and Shore 1999)。第二步：Sir4-Sir2 和 Sir4-Sir3 发生强相互作用，于是沿着 TG 重复形成了 Sir 复合体。复合体的形成刺激了 Sir2 的 NAD 依赖的组蛋白去乙酰化活性，然后 Sir2 催化了邻近组蛋白 H4K16 位点的去乙酰化。第三步：Sir 复合体沿着核小体蔓延，或许利用了 NAD 水解产生的 O-乙酰基核糖（Liou et al. 2005)。Sir3 和 Sir4 结合非乙酰化的组蛋白 H4 尾部。虽然非乙酰化的组蛋白 H3 尾部也可以结合 Sir3 和 Sir4，但本图中未显示。第四步：沉默的染色体在 M 期末装配完成，产生了一个不易被接近的结构。这可能必须有更高程度的折叠并被置于核膜处。

在异染色质成核事件中有协同作用（Stavenhagen and Zakian 1994）。值得一提的是在散布在酵母基因组中的启动子上，Rap1 可以作为基本转录因子激活很多基因的表达，特别是那些编码核糖体蛋白的基因。现在还不知道为什么 Rap1 招募激活子到这些启动子上却不引起异染色质形成。

## 6. Sir2 催化的组蛋白去乙酰化为 SIR 复合体蔓延提供结合位点

SIR 蛋白的分子间相互作用已经被明确，Sir4 对他们的聚合起关键作用。Sir4 在体外与 Sir2 有强烈的相互作用。Sir4 也与 Sir3 有独立的相互作用，但是 Sir3 和 Sir2 的相互作用很微弱（Moazed et al. 1997；Strahl-Bolsinger et al. 1997；Hoppe et al. 2002）。Sir3 和 Sir4 可以形成同源二聚体（Moretti et al. 1994）。尽管如此，当在昆虫细胞中共表达时，可以分离出一个稳定的 360 kD 复合体，其中 Sir2、Sir3 和 Sir4 的比例为 1∶1∶1（Cubizolles et al. 2006）。与 Sir2、Sir3 和 Sir4 形成一个功能的异源三聚体复合体相一致的是，染色质免疫共沉淀也表明这三种 SIR 蛋白沿着异染色质区域有相同的分布。但是很明显，Sir3 在这个过程中有着特殊的作用，因为过表达 Sir3 可以扩大沉默结构域的范围，伴随着 Sir3 从正常的 3 kb 蔓延到 15 kb（Renauld et al. 1993；Hecht et al. 1996）。不平衡表达 Sir2 或 Sir4，或仅表达这两个蛋白的一部分结构域，都产生恰好相反的效果：过表达 Sir2 或者 Sir4 破坏端粒位置效应，而同时表达 Sir3 和 Sir4 抵消掉这种不平衡时则可以恢复沉默（Maillet et al. 1996）。这揭示了 SIR 复合体中三种成分的比例对于它的抑制功能十分重要，就如果蝇中 Polycomb 复合体一样。过表达的 Sir3 有沿着染色质蔓延的独特能力，与此相一致的是 Sir3 在体外可以形成稳定的多聚体（Liou et al. 2005）。

核小体上去乙酰化的组蛋白 H3 和 H4 的 N 端构成了 SIR 复合体蔓延的平台

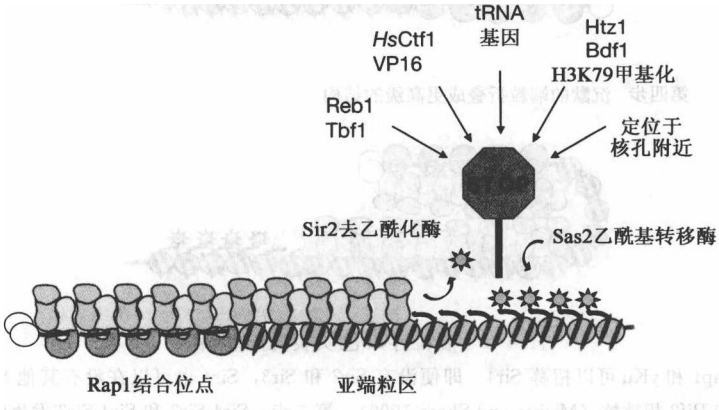



图 4-7 酿酒酵母中的异染色质边界

邻近的常染色质上由 Sas2 催化的 H4K16 乙酰化限制了由 Sir2 催化的 H4K16 去乙酰化，导致的异染色质蔓延，因为其阻碍了 Sir3 的结合。邻近常染色质上 H3K79 甲基化同样影响了异染色质的蔓延。另外，Reb1、Tbf1、哺乳动物或者病毒的 Ctf1 或 VP16 定位于核孔附近、tRNA 基因的存在可能也调节了边界的功能。可以推断，其中某些成分也通过招募组蛋白去乙酰化酶来发挥作用。

(Braunstein et al. 1996; Suka et al. 2001)。SIR 复合体与组蛋白的结合方式帮助解释了蔓延是如何发生的（图 4-7)。在体内以及体外 Sir3 和 Sir4 可以与未乙酰化的 H3 和 H4 的氨基端相结合（Hecht et al. 1995，1996)，这种结合不需要 H2A 和 H2B 的尾部。对此最重要的组蛋白部位是 H4 第 16～29 位氨基酸残基，第 16 位赖氨酸必须去乙酰化或带正电荷才能与 Sir3 结合（Johnson et al. 1990，1992)。不像其他乙酰化位点的突变，H4K16 即便发生一个保守的突变都会完全破坏端粒区域的沉默。组蛋白 H3/H4 的尾部，特别是 H4 的 16～24 位残基，在体外也促进核小体阵列紧密排列，H4K16 的乙酰化状态似乎在这个过程中调节核小体纤维的高度有序折叠。然而 H4K16 去乙酰化在体内是如何被调控的呢?

## 7. Sir2 催化组蛋白 H4 第 16 位赖氨酸的去乙酰化

Sir2 是一个 NAD 依赖的组蛋白去乙酰化酶，与 Sir4 结合可增强其活性。Sir2 利用 ADP-核糖基转移酶活性使 NAD 转变成 O-乙酰基 ADP 核糖来实现去乙酰化（Tanner et al. 2000)。既然一个带正电荷的 H4K16 对异染色质形成有关键作用，Sir2 能在体内和体外去乙酰化 H4K16 就显得尤为重要，虽然它也能去乙酰化 H4 氨基端的其他赖氨酸以及 H3K9/H3K14（Imai et al. 2000; Suka et al. 2002; Cubizolles et al. 2006)。所有这些靶位点都存在于 H3 和 H4 与沉默相关的结构域内。有趣的是，O-乙酰基 ADP 核糖不但促进 Sir3 的多聚化，而且促进 Sir3 与 Sir4-Sir2 在体外的相互作用（Liou et al. 2005)。这些数据证明 Sir2 介导的组蛋白去乙酰化可促进 SIR 复合体的形成，并在临近的核小体上为 SIR 复合体提供了去乙酰化了的结合位点。

在图 4-6 中我们总结了在端粒和 HM 位点处异染色质的起始与蔓延。在端粒处，Rap1 招募 Sir4，Sir4 招募 Sir2，Sir2 催化组蛋白 H4 和 H3 氨基端的去乙酰化。Sir4 也招募 Sir3。组蛋白尾部的去乙酰化产生了 Sir3/Sir4 的结合位点，然后导致 SIR2-3-4 复合体的结合。Sir3 与 Sir4 的相互作用，以及它们与组蛋白氨基端的相互作用是为了在核小体纤维上稳定 SIR 复合体，以使它能沿着组蛋白尾部蔓延。最后，染色质纤维（下面会讨论到）的折叠可以稳定抑制状态。在 HM 位点处发生的大部分事件与此相似，虽然 Sir4 最初是通过 Rap1、Abf1、ORC 和 Sir1 招募的。那么这种蔓延是如何被终止的呢?

## 8. 常染色质上的组蛋白乙酰化阻止 SIR 复合体的蔓延

既然 Sir2 催化的 H4K16 去乙酰化对于异染色质的形成起着关键的作用，所以不难理解这个位点的修饰在阻止异染色质的传播中也扮演一个重要角色。有趣的是，对于常染色质上单乙酰化的 H4，所有的乙酰化位点中只有 K16 是被修饰的（Clarke et al. 1993)。其中一个负责酵母亚端粒区 H4K16 乙酰化的酶是 Sas2，它是乙酰基转移酶（HAT）中高保守的 MYST 家族的成员。如果 Sas2 被敲除从而阻止 H4K16 的乙酰化，或者 H4K16 被突变成精氨酸来模仿去乙酰化的状态时，与野生型细胞相比，染色体 VI

的右侧端粒上 SIR 复合体会以较低的水平蔓延到五倍以上的距离。这至少部分的表明，在亚端粒区异染色质的蔓延由 Sir2 和 Sas2 间的拮抗活性来控制（图 4-7）（Kimura et al. 2002；Suka et al. 2002）。

与端粒区相比，也许在 HM 位点两侧限制沉默染色质的蔓延更为重要，因为沿着Ⅲ号染色体臂分布着许多对生长很重要的基因，而且沉默子作用是双向的，可以沉默两侧的 DNA 序列。阻止从 HMR 到端粒方向的沉默蔓延的一条分界线是一个 tRNA 基因（Donze and Kamakaka 2001）。这种分界线功能似乎需要与这个区域的转录或转录潜能相关的 HAT 活性。重要的是其中的一个 HAT 是 Sas2，虽然 Gcn5（组蛋白 H3 的 HAT）也可以促进此 tRNA 基因的分界线功能。这表明转录激活因子可以通过招募 HAT 来限制 SIR 复合体的蔓延。相应地，分界线活性也可以归功于转录因子 Reb1、Tbf1、哺乳动物的转录因子 CTCF 以及含有酸性激活结构域的 VP16（Fourel et al. 1999，2001）。所有的这些因子也可以促进组蛋白的高乙酰化，从而通过破坏 SIR 复合体与核小体的结合来阻止其蔓延（图 4-7）。

最后，在上述机制之外，有报道表明存在于常染色质上的组蛋白变体 H2A. Z、RNA 聚合酶结合因子 Bdf1（Meneghini et al. 2003）、H3K79 的甲基化（van Leeuwen et al. 2002）以及 DNA 定位于核孔附近都有助于限制沉默染色质的蔓延。虽然这些因素影响异染色质蔓延的机制还不清楚，但有趣的是一些活跃基因结合于核孔周边（Ishii et al. 2002；Brickner and Walter 2004）。因此，酵母中边界因子的共同特点是作为强转录激活因子或者核小体重塑因子，它们直接或者间接地破坏组蛋白与异染色质蛋白间的相互作用。

## 9. 端粒成环

一些证据表明长距离的相互作用可以使染色体末端折回，绕开亚端粒边界元件，从而稳定地抑制亚端粒区的基因（图 4-5、图 4-6）。例如，尽管 Rap1 结合位点仅仅存在于端粒末端约 300 bp 的 TG 重复 DNA，染色质免疫共沉淀分析表明 Rap1 可以与距 TG 重复 3 kb 之外的核小体相结合（Strahl-Bolsinger et al. 1997）。相似的，从端粒末端起 3 kb均有 yKu 的结合（Martin et al. 1999）。更重要的是，当 SIR 基因突变破坏沉默后，Rap1 和 yKu 仅仅只从更靠近内部的序列上消失，而仍存在于末端的 TG 重复序列（Martin et al. 1999）。染色质已被剪切成小于 500 bp 的片段，但结合有 Rap1 和/或 yKu 的异染色质仍有 3 kb，为了解释这一点，有人提出端粒折回使得结合于 TG 序列的 Rap1 和 yKu 与分布于染色体上的 SIR 蛋白反式结合（图 4-5、图 4-6）。这种结构可能有助于端粒结合蛋白的“帽子”功能。

支持端粒成环的证据来自于 Bruin 等的工作（2001），他们利用了转录激活因子，如 Gal4 不能在它所要激活的基因的下游发挥作用的特点。他们把 Gal4 上游激活序列（upstream activating sequence，UAS）元件构建于报告基因的下游，然后分别插入染色体内部以及靠近端粒的位置。当在一个内部位点时，转录不能被半乳糖诱导。然而，处在亚端粒区时，Gal4 UAS 可以以一种依赖于 Sir3 的方式在启动子下游 1.9 kb 的位

置激活这个启动子。研究人员认为，在 Sir3 存在的情况下端粒末端可以折回，从而使 Gal4 UAS 接近转录起始位点（de Bruin et al. 2001）。

## 10. 端粒成环介导的天然亚端粒元件处的抑制不连续性

我们在此已提出了一个简化了的观点，即从端粒末端 Rap1 结合位点开始出现连续的沉默染色质，但是在天然端粒中情况要复杂许多，主要是因为在亚端粒重复序列中存在天然的边界元件。一般来说，当报告基因被插入的时候，亚端粒重复序列 X 和 Y′都被缺失，使报告基因和特定序列紧邻于 TG 重复。而另一方面，所有天然的端粒都含有一个核心的亚端粒重复元件 X，其位于 TG 重复与最靠近端粒的基因之间，50%～70%天然的端粒也至少含有一个拷贝的较大的亚端粒元件 Y′（图 4-8）。X 和 Y′元件都含有转录调节子 Tbf1 和 Reb1 的结合位点，被认为可以减少沉默染色质的蔓延（Fourel et al. 1999）。然而，X 元件也含有 ORC 和 Abf1 的结合位点，它们则有相反的效果：重起始或促进对元件靠近端粒一边的报告基因的抑制。结果是天然的端粒沉默中间有不连续性，这与图 4-6 中描述的连续的蔓延模型有所不同。为了解释这一点，Pryde 和 Lou-

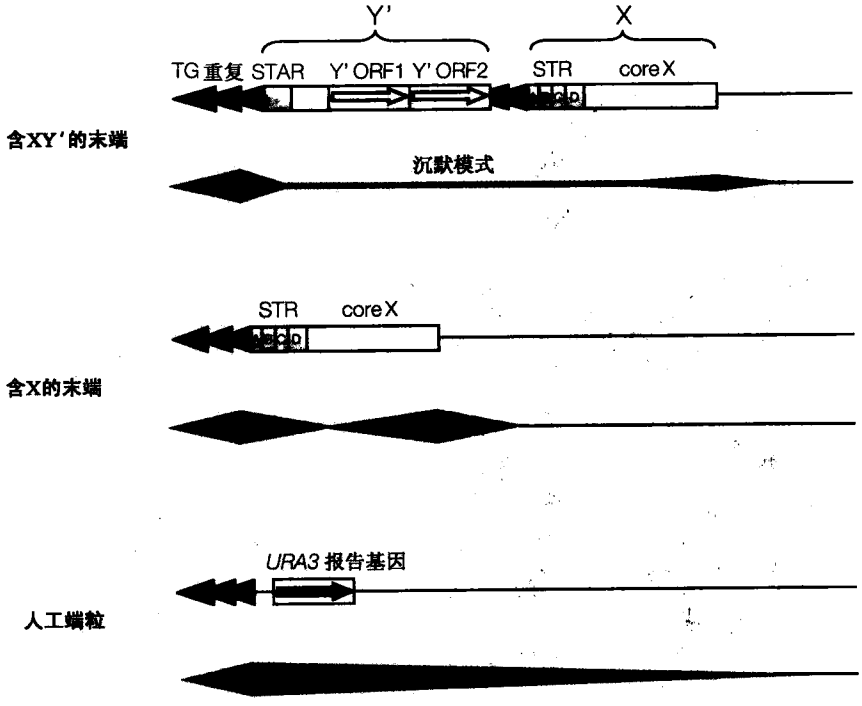



图 4-8　天然端粒的组织结构及其沉默模式

图中所示的是亚端粒元件及其主要的蛋白结合位点。端粒一般可以分成两类：一种含有 X，一种含有 X 和 Y′。STAR 和 STR 元件阻止了抑制状态的蔓延，并且在 X 或 Y′内留下低抑制水平的区域。在人工截短的端粒中却并非如此，那里的情况是从 TG 重复延伸出 3～4 kb 的成梯度减弱抑制区域。天然端粒被认为形成类似图 4-6 所示的环，抑制的区域是彼此接触的，在相接触区域之间的是非抑制的染色质（摘自 Pryde and Louis 1999）。

is（1999）提出是端粒折回使一个非抑制性的染色质区域插入到两个抑制性区域之间，在不影响从 TG 重复序列开始的成核和蔓延的情况下，导致在沉默区域中出现不连续性。

## 11. 端粒间的反式相互作用和异染色质在核周的附着

异染色质最普遍且保守的特征之一是其出现于离散的核内亚区室。在酿酒酵母中也是如此，端粒在间期聚集成簇，与核周保持紧密结合。这种簇最早在以下实验中被发现：在一个弥散的核背景中用免疫染色观察 Rap1 和 SIR 蛋白，可以发现明显的聚集（图 4-9）。通过突变 H4K16 或者干扰 Rap1 或 yKu 的功能来破坏沉默，导致 SIR 蛋白从这些簇上散布到各处（Hecht et al. 1995；Laroche et al. 1998）。后来发现不止端粒，HML 和 HMR 位点也和核膜紧密结合。这种结合由若干冗余的通路所介导，或者依赖于端粒结合因子 yKu，或者依赖于沉默染色质的形成本身（Hediger et al. 2002）。在沉默染色质内，这种锚定的功能由 Sir4 的亚结构域来完成，其可以和位于核膜的蛋白 Esc1 相结合（enhances silent chromatin 1；Taddei et al. 2004）。Sir4-Esc1 的相互作用使被 SIR 蛋白抑制的染色质结构域定位于核膜，且位于核孔之外的位点。甚至在 yKu 锚定通路不存在的情况下，只要抑制状态得到维持，Sir4-Esc1 的相互作用也可以保持

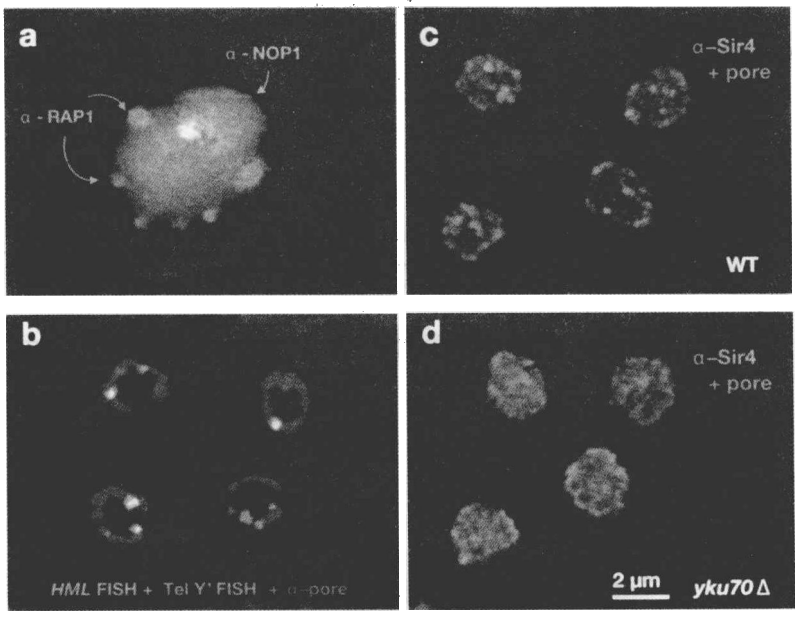



图 4-9 SIR 蛋白和 Rap1 被发现呈点状集中分布于核周

a，Rap1（绿色）显示的七个簇代表了二倍体细胞中的 64 个端粒。他们或位于核周或者临近核仁（蓝色，抗 Nop1）。红色为 DNA。b，用荧光原位杂交（FISH）显示端粒 DNA，绿色为 HML。两者约有 70%是共定位的，并且都靠近核膜（蓝色）（Heun et al. 2001）。c，Sir4（绿色）在核膜（Mab414 染色为红色）附近的点状聚集分布。这种分布在缺失 yKu70 的菌株中消失，对应着端粒沉默状态的丢失（Laroche et al. 1998）。

端粒与核膜的联系。通过重组的方法将沉默的染色质环切出，从而与邻近的端粒分开后，仍可以以 SIR 蛋白依赖的方式与核周保持结合（Gartenberg et al. 2004）。

最初把端粒招募到核膜的或许是 yKu，因为在沉默不发生的情况下，这一功能也可以实现。这种锚定与端粒间的反式相互作用一起导致了一个核内亚区室的形成，进而依次结合 SIR 蛋白。这种亚区室对于在端粒处形成一种沉默梯度是十分重要的。甚至，当两边有沉默子的 HM 位点被插入到端粒区域（Thompson et al. 1994；Maillet et al. 1996）或者通过跨膜因子人为定位到核膜附近时（Andrulis et al. 1998），都能更有效地导致沉默。重要的是，当 Sir3 和 Sir4 不再呈点状聚集分布或被过表达时，这种因靠近端粒导致的而促进抑制的能力消失了（Maillet et al. 1996；Marcand et al. 1996）。这证明了 SIR 蛋白浓度梯度是端粒簇促进抑制的重要特征。最后，有人指出，以有限浓度存在的通用抑制因子的聚集有助于使细胞实现沉默的表观遗传状态的继承，描述见图 4-10。简而言之，这个模型认为，如果 DNA 在一个沉默因子富集的亚区室内复制，那么将更容易形成异染色质。

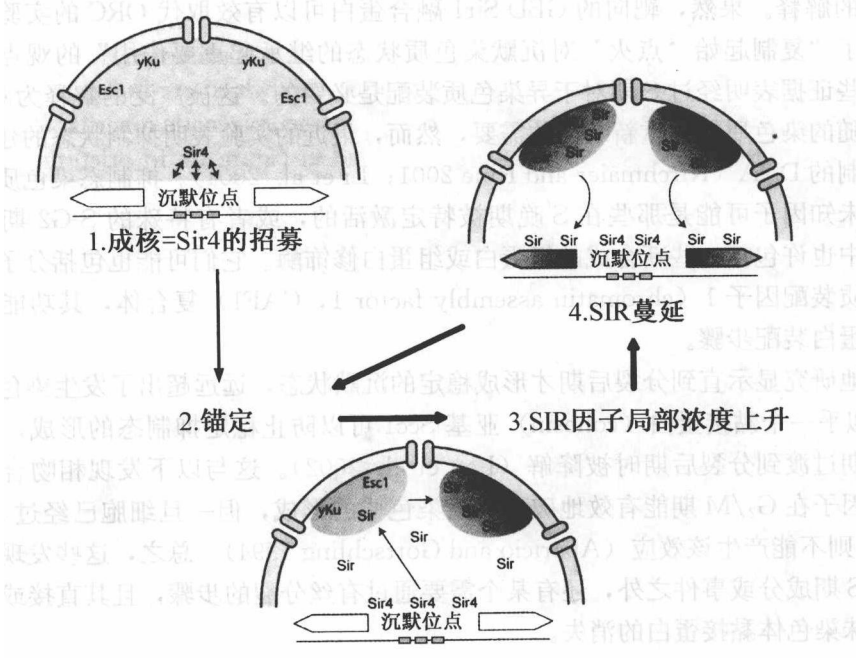



图 4-10　沉默亚区室的自发形成

图中所示的是一个核内亚区室形成的简单模型。（1）最初，可结合 Sir4 的 DNA 结合蛋白（包括 Rap1、ORC、Abf1 和 yKu）把 Sir4 招募到成核中心。（2）通过两个 Sir4 锚定途径（yKu 或 Esc1）之一，Sir4 把其所结合的区域带至核周。（3）在核膜，高浓度的 SIR 蛋白有助于沉默复合体的装配和蔓延。（4）沉默区域与核周的结合提高了 SIR 蛋白在局部区域内的浓度，进而增强了沉默该区域内其他位点的能力。重要的是，就像 Sir4 招募沉默子序列一样，可与端粒结合的 yKu 能独立地把端粒招募到核周。

## 12. 表观遗传状态的继承

异染色质的一个普遍特性是它的沉默状态可以从一代传递到下一代。这需要在DNA复制后的子链上很快发生异染色质结构的重新装配。关于细胞周期如何影响染色质状态的建立和继承方面的先驱工作由Miller和Nasmyth（1984）进行，他们用一个温度敏感的*sir3*$^{ts}$突变体来研究沉默的起始与丢失。从许可温度到非许可温度的转换引起沉默的迅速丢失，这表明维持抑制状态需要Sir3。然而，在相反的实验中，非允许温度到允许温度的转换并没有立即导致沉默的恢复：需要通过一轮细胞周期。他们得出结论抑制的染色质的建立和遗传需要S期的某个事件。后来的研究表明还需要G2/M期的事件（Lau et al. 2002）。

最初，人们认为沉默子连锁的ARS的复制起始“点火”可能是沉默染色质状态建立或继承中的一个关键事件，但是由于复制起始并未在HML位点被检测到，这似乎不是正确的解释。果然，靶向的GBD-Sir1融合蛋白可以有效取代ORC的实验使人们彻底摒弃了“复制起始‘点火’对沉默染色质状态的继承起重要作用”的观点。另一方面，一些证据表明经过S期对于异染色质装配是必要的。这被广泛的解释为对DNA复制及伴随的染色质结构重新装配的需要，然而，最近的实验表明抑制状态的建立能发生于不复制的DNA（Kirchmaier and Rine 2001；Li et al. 2001）。抑制态染色质建立所需的一些未知因子可能是那些在S晚期被特定激活的，或者有特殊的S-G2期功能的蛋白。其中也许包括一些新合成的组蛋白或组蛋白修饰酶。它们可能也包括分子伴侣，例如染色质装配因子1（chromatin assembly factor 1，CAF1）复合体，其功能为确保重要的组蛋白装配步骤。

其他研究显示直到分裂后期才形成稳定的沉默状态，远远超出了发生染色体装配的S期。似乎一个黏接蛋白（cohesin）亚基Scc1可以防止稳定抑制态的形成，直到其在分裂中期过渡到分裂后期时被降解（Lau et al. 2002）。这与以下发现相吻合，靶向导入转录因子在$G_2$/M期能有效地破坏沉默染色质的形成，但一旦细胞已经过M期进入G1期，则不能产生该效应（Aparicio and Gottschling 1994）。总之，这些发现认为，在关键的S期成分或事件之外，还有某个需要通过有丝分裂的步骤，且其直接或间接的依赖于姐妹染色体黏接蛋白的消失。

## 13. 衰老与Sir2：因rDNA重复序列的不稳定性而相联系

在果蝇中，高度活跃的rDNA重复序列与着丝粒异染色质相邻，在很多高等真核物种中，核仁和压缩态的异染色质在空间上相邻。因此，非常重要的是，在酵母中，Sir2以不依赖其他SIR蛋白的方式，与高转录活性的rDNA重复有遗传上和物理上的相互作用（Gotta et al. 1997）。确实，Sir2抑制rDNA的重组，也抑制插入到rDNA阵列中的依赖RNA聚合酶II的报告基因的表达。鉴于rDNA重复序列首尾相连重复排列，易发生不对称重组事件，导致rDNA阵列的减少或者增加。这种不稳定性与观察

到的染色体外 rDNA 环的累积现象相关（图 4-11）（Sinclair and Guarente 1997）。抑制插入到 rDNA 中的报告基因及防止可导致 rDNA 重复丢失的异常重组，都需要 Sir2；这可能是通过置入核小体（Fritze et al. 1997）或使姐妹染色单体中重复序列相匹配来阻止不对称重组事件的发生（Kobayashi et al. 2004）。

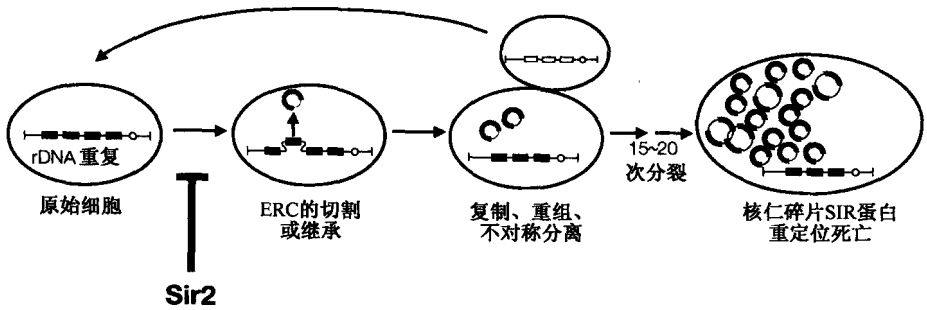



图 4-11　rDNA 重组导致酵母细胞衰老

rDNA 由 140～200 个 9.1 kb 的单元排列而成（红色方框）。他们编码 18S、5.8S、25S 和 5SrRNA，在 5S 基因下游和 18S 基因中存在两个 Sir2 应答元件。rDNA 重复在衰老的酵母细胞中易被切割出，切割而成的环可以在母细胞中累积（Kaeberlein et al. 1999）。这与过早衰老相关并且能被 Sir2 抑制，Sir2 能帮助抑制不对称重组与 rDNA 环的切割。

酵母 Sir2 缺失最令人惊讶的表型是寿命的缩短，而在酵母中寿命的长短与端粒抑制的丢失或者 TG 重复区的长度无直接联系。事实上，Sir2 缺陷的酵母寿命变短的表型意味着这些细胞平均分裂少于 12 次，而非野生型细胞的 20～25 次（Kaeberlein et al. 1999）。现在人们已经确信因不对称重组事件所产生的染色体外的 rDNA 重复序列环（extrachromosomal rDNA repeat circles，ERC），及其在母细胞中的积累与酵母的衰老相关（图 4-11）。重要的是，生命期不仅因缺失 Sir2 而变短，而且还可以通过过表达 Sir2 而延长，过表达 Sir2 导致结合到 rDNA 上的 Sir2 增多。另外有一些突变体可以降低 rDNA 切割效率（Defossez et al. 1999），例如，复制叉障碍蛋白 Fob1 的突变可以延长酵母生命期，正如人为产生 ERC 可以导致细胞的衰老（Sinclair and Guarente 1997）。因此，rDNA 不稳定性显然与酵母衰老有关，尽管其作用可能是间接的。有一种模型认为高水平的 ERC 占用了 DNA 修复或复制相关蛋白，导致基因组损伤的增加，或降低了基因组其余部分的复制效率。

因为 Sir2 是一个 NAD 依赖的去乙酰化酶，而 NAD 水平又是新陈代谢的指标，有人认为 Sir2 调节寿命的功能可能与卡路里限制对寿命的延长作用有关，后者是许多物种中保守存在的一条延缓衰老的通路。尽管这一观点得到了以下实验的支持，即卡路里限制可以在酵母、果蝇和哺乳动物中提高 Sir2 的活性。但研究表明酵母寿命在低葡萄糖条件（卡路里限制）下的延长，并不依赖于 Sir2，且与 Sir2 作用有累加效应。因此，Sir2 和卡路里限制通过不同的通路来延长寿命。

切割出的 rDNA 环的累积尚未在其他任何物种中被发现，但人们认为在秀丽线虫和啮齿类中其他形式的基因组不稳定与生命期变短相关。与酿酒酵母中缺失 Sir2 导致姐妹染色体间的不对称重组类似，哺乳动物中端粒丢失了异染色质状态后会导致染色体

末端之间的融合，从而限制细胞的分裂潜能。虽然我们还不知道哺乳动物的 Sir2 是否影响这些机制，但似乎基因组不稳定性是衰老中的一个共同因素，而异染色质结构的丢失很可能专一性地影响了这些事件。

## 14. 总　结

遗传学、生物化学和细胞生物学技术共同被用于研究酿酒酵母中异染色质介导的基因沉默的基本规律。这些规律包括异染色质形成的起始、蔓延以及对蔓延的阻止；异染色质因子的平衡及其在核内亚区室中的分布；异染色质环化以及细胞周期在异染色质形成中的作用。此外，在体外重构酵母异染色质的系统已经被建立。这些体内和体外的研究为我们理解所有真核生物中从染色质纤维到异染色质的组装提供了良好的基础。

马　奔　译
杨　鹏　校
朱　冰　审校

## 参考文献

Andrulis E.D., Neiman A.M., Zappulla D.C., and Sternglanz R. 1998. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394:** 592–595.

Aparicio O.M. and Gottschling D.E. 1994. Overcoming telomeric silencing: A trans-activator competes to establish gene expression in a cell cycle-dependent way. *Genes Dev.* **8:** 1133–1146.

Aparicio O.M., Billington B.L., and Gottschling D.E. 1991. Modifiers of position effect are shared between telomeric and silent mating-type loci in *S. cerevisiae*. *Cell* **66:** 1279–1287.

Boscheron C., Maillet L., Marcand S., Tsai-Pflugfelder M., Gasser S.M., and Gilson E. 1996. Cooperation at a distance between silencers and proto-silencers at the yeast *HML* locus. *EMBO J.* **15:** 2184–2195.

Brand A.H., Micklem G., and Nasmyth K. 1987. A yeast silencer contains sequences that can promote autonomous plasmid replication and transcriptional activation. *Cell* **51:** 709–719.

Braunstein M., Sobel R.E., Allis C.D., Turner B.M., and Broach J.R. 1996. Efficient transcriptional silencing in *Saccharomyces cerevisiae* requires a heterochromatin histone acetylation pattern. *Mol. Cell. Biol.* **16:** 4349–4356.

Brickner J.H. and Walter P. 2004. Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS Biol.* **2:** e342.

Chen L. and Widom J. 2005. Mechanism of transcriptional silencing in yeast. *Cell* **120:** 37–48.

Cheng T.H. and Gartenberg M.R. 2000. Yeast heterochromatin is a dynamic structure that requires silencers continuously. *Genes Dev.* **14:** 452–463.

Chien C.T., Buck S., Sternglanz R., and Shore D. 1993. Targeting of SIR1 protein establishes transcriptional silencing at HM loci and telomeres in yeast. *Cell* **75:** 531–541.

Chopra V.S. and Mishra R.K. 2005. To SIR with Polycomb: Linking silencing mechanisms. *Bioessays* **27:** 119–121.

Clarke D.J., O'Neill L.P., and Turner B.M. 1993. Selective use of H4 acetylation sites in the yeast *S. cerevisiae*. *Biochem. J.* **294:** 557–561.

Cubizolles F., Martino F., Perrod S., and Gasser S.M. 2006. A homotrimer-heterotrimer switch in Sir2 structure differentiates rDNA and telomeric silencing. *Mol. Cell* **21:** 825–836.

de Bruin D., Zaman Z., Liberatore R.A., and Ptashne M. 2001. Telomere looping permits gene activation by a downstream UAS in yeast. *Nature* **409:** 109–113.

Defossez P.A., Prusty R., Kaeberlein M., Lin S.J., Ferrigno P., Silver P.A., Keil R.L., and Guarente L. 1999. Elimination of replication block protein Fob1 extends the life span of yeast mother cells. *Mol. Cell* **3:** 447–455.

Donze D. and Kamakaka R.T. 2001. RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in *S. cerevisiae*. *EMBO J.* **20:** 520–531.

Fourel G., Revardel E., Koering C.E., and Gilson E. 1999. Cohabitation of insulators and silencing elements in yeast subtelomeric regions. *EMBO J.* **18:** 2522–2537.

Fourel G., Boscheron C., Revardel E., Lebrun E., Hu Y.F., Simmen K.C., Muller K., Li R., Mermod N., and Gilson E. 2001. An activation-independent role of transcription factors in insulator function. *EMBO Rep.* **2:** 124–132.

Fritze C.E., Verschueren K., Strich R., and Easton Esposito R. 1997. Direct evidence for SIR2 modulation of chromatin structure in yeast rDNA. *EMBO J.* **16:** 6495–6509.

Frye R.A. 2000. Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochem. Biophys. Res. Comm.* **273:** 793–798.

Gartenberg M.R., Neumann F.R., Laroche T., Baszczyk M., and Gasser S.M. 2004. Sir-mediated repression can occur independently of chromosomal and subnuclear contexts. *Cell* **119:** 955–967.

Gotta M., Strahl-Bolsinger S., Renauld H., Laroche T., Kennedy B.K., Grunstein M., and Gasser S.M. 1997. Localization of Sir2p: The nucleolus as a compartment for silent information regulators. *EMBO J.* **16:** 3243–5503.

Gottlieb S. and Esposito R.E. 1989. A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA. *Cell* **56:** 771–776.

Gottschling D.E. 1992. Telomere-proximal DNA in *Saccharomyces cerevisiae* is refractory to methyltransferase activity in vivo. *Proc. Natl. Acad. Sci.* **89:** 4062–4065.

Gottschling D.E., Aparicio O.M., Billington B.L., and Zakian V.A. 1990. Position effect at *S. cerevisiae* telomeres: Reversible repression of PolII transcription. *Cell* **63:** 751–762.

Hecht A., Strahl-Bolsinger S., and Grunstein M. 1996. Spreading of

transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* **383:** 92–96.

Hecht A., Laroche T., Strahl-Bolsinger S., Gasser S.M., and Grunstein M. 1995. Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: A molecular model for the formation of heterochromatin in yeast. *Cell* **80:** 583–592.

Hediger F., Neumann F.R., Van Houwe G., Dubrana K., and Gasser S.M. 2002. Live imaging of telomeres: yKu and Sir proteins define redundant telomere-anchoring pathways in yeast. *Curr. Biol.* **12:** 2076–2089.

Heun P., Laroche T., Raghuraman M.K., and Gasser S.M. 2001. The positioning and dynamics of origins of replication in the budding yeast nucleus. *J. Cell Biol.* **152:** 385–400.

Hoppe G.J., Tanny J.C., Rudner A.D., Gerber S.A., Danaie S., Gygi S.P., and Moazed D. 2002. Steps in assembly of silent chromatin in yeast: Sir3-independent binding of a Sir2/Sir4 complex to silencers and role for Sir2-dependent deacetylation. *Mol. Cell. Biol.* **12:** 4167–4180.

Imai S.I., Armstrong C., Kaeberlein M., and Guarente L. 2000. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403:** 795–800.

Ishii K., Arib G., Lin C., Van Houwe G., and Laemmli U.K. 2002. Chromatin boundaries in budding yeast: The nuclear pore connection. *Cell* **109:** 551–562.

Johnson L.M., Fisher-Adams G., and Grunstein M. 1992. Identification of a non-basic domain in the histone H4 N-terminus required for repression of the yeast silent mating loci. *EMBO J.* **11:** 2201–2209.

Johnson L.M., Kayne P.S., Kahn E.S., and Grunstein M. 1990. Genetic evidence for an interaction between SIR3 and histone H4 in the repression of the silent mating loci in *S. cerevisiae*. *Proc. Natl. Acad. Sci.* **87:** 6286–6290.

Kaeberlein M., McVey M., and Guarente L. 1999. The SIR2/3/4 complex and SIR2 alone promote longevity in *S. cerevisiae* by two different mechanisms. *Genes Dev.* **13:** 2570–2580.

Kaeberlein M., Kirkland K.T., Fields S., and Kennedy B.K. 2004. Sir2-independent life span extension by calorie restriction in yeast. *PLoS Biol.* **2:** 296–307.

Kayne P.S., Kim U.J., Han M., Mullen J.R., Yoshizaki F., and Grunstein M. 1988. Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell* **55:** 27–39.

Kimura A., Umehara T., and Horikoshi M. 2002. Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing. *Nat. Genet.* **3:** 370–377.

Kirchmaier A.L. and Rine J. 2001. DNA replication-independent silencing in *S. cerevisiae*. *Science* **291:** 646–650.

Kobayashi T., Horiuchi T., Tongaonkar P., Vu L., and Nomura M. 2004. Sir2 regulates recombination between different rDNA repeats, but not recombination within individual rRNA genes in yeast. *Cell* **117:** 441–453.

Laroche T., Martin S.G., Gotta M., Gorham H.C., Pryde F.E., Louis E.J., and Gasser S.M. 1998. Mutation of yeast Ku genes disrupts the subnuclear organization of telomeres. *Curr. Biol.* **8:** 653–656.

Lau A., Blitzblau H., and Bell S.P. 2002. Cell-cycle control of the establishment of mating-type silencing in *S. cerevisiae*. *Genes Dev.* **16:** 2935-2945.

Li Y.C., Cheng T.H., and Gartenberg M.R. 2001. Establishment of transcriptional silencing in the absence of DNA replication. *Science* **291:** 650–653.

Liou G.G., Tanny J.C., Kruger R.G., Walz T., and Moazed D. 2005. Assembly of the SIR complex and its regulation by *O*-acetyl-ADP-ribose, a product of NAD-dependent histone deacetylation. *Cell* **121:** 515–527.

Loo S. and Rine J. 1994. Silencers and domains of generalized repression. *Science* **264:** 1768–1771.

Luo K., Vega-Palas M.A., and Grunstein M. 2002. Rap1-Sir4 binding independent of other Sir, yKu, or histone interactions initiates the assembly of telomeric heterochromatin in yeast. *Genes Dev.* **12:** 1528–1539.

Maillet L., Boscheron C., Gotta M., Marcand S., Gilson E., and Gasser S.M. 1996. Evidence for silencing compartments within the yeast nucleus: A role for telomere proximity and Sir protein concentration in silencer-mediated repression. *Genes Dev.* **10:** 1796–1811.

Marcand S., Gilson E., and Shore D. 1997. A protein-counting mechanism for telomere length regulation in yeast. *Science* **275:** 986–990.

Marcand S., Buck S.W., Moretti P., Gilson E., and Shore D. 1996. Silencing of genes at nontelomeric sites in yeast is controlled by sequestration of silencing factors at telomeres by Rap1 protein. *Genes Dev.* **10:** 1297–1309.

Martin S.G., Laroche T., Suka N., Grunstein M., and Gasser S.M. 1999. Relocalization of telomeric Ku and SIR proteins in response to DNA strand breaks in yeast. *Cell* **97:** 621–633.

Meneghini M.D., Wu M., and Madhani H.D. 2003. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* **112:** 725–736.

Miller A.M. and Nasmyth K.A. 1984. Role of DNA replication in the repression of silent mating type loci in yeast. *Nature* **312:** 247–251.

Mishra K. and Shore D. 1999. Yeast Ku protein plays a direct role in telomeric silencing and counteracts inhibition by Rif proteins. *Curr. Biol.* **9:** 1123–1126.

Moazed D., Kistler A., Axelrod A., Rine J., and Johnson A.D. 1997. Silent information regulator protein complexes in *S. cerevisiae*: A SIR2/SIR4 complex and evidence for a regulatory domain in SIR4 that inhibits its interaction with SIR3. *Proc. Natl. Acad. Sci.* **94:** 2186–2191.

Moretti P., Freeman K., Coodly L., and Shore D. 1994. Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes Dev.* **8:** 2257–2269.

Palladino F., Laroche T., Gilson E., Axelrod A., Pillus L., and Gasser S.M. 1993. SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. *Cell* **75:** 542–555.

Pillus L. and Rine J. 1989. Epigenetic inheritance of transcriptional states in *S. cerevisiae*. *Cell* **59:** 637–647.

Pryde F.E. and Louis E.J. 1999. Limitations of silencing at native yeast telomeres. *EMBO J.* **18:** 2538–2550.

Renauld H., Aparicio O.M., Zierath P.D., Billington B.L., Chhablani S.K., and Gottschling D.E. 1993. Silent domains are assembled continuously from the telomere and are defined by promoter distance and strength, and by SIR3 dosage. *Genes Dev.* **7:** 1133–1145.

Rusche L.N., Kirchmaier A.L., and Rine J. 2002. Ordered nucleation and spreading of silenced chromatin in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **7:** 2207–2222.

———. 2003. The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* **72:** 481–516.

Sekinger E.A. and Gross D.S. 1999. SIR repression of a yeast heat shock gene: UAS and TATA footprints persist within heterochromatin. *EMBO J.* **18:** 7041–7055.

Sinclair D.A. and Guarente L. 1997. Extrachromosomal rDNA circles—A cause of aging in yeast. *Cell* **91:** 1033–1042.

Stavenhagen J.B. and Zakian V.A. 1994. Internal tracts of telomeric DNA act as silencers in *Saccharomyces cerevisiae*. *Genes Dev.* **8:** 1411–1422.

Strahl-Bolsinger S., Hecht A., Luo K., and Grunstein M. 1997. SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.* **11:** 83–93.

Suka N., Luo K., and Grunstein M. 2002. Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat. Genet.* **3:** 378–383.

Suka N., Suka Y., Carmen A.A., Wu J., and Grunstein M. 2001. Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Mol. Cell* **8:** 473–479.

Taddei A., Hediger F., Neumann F.R., Bauer C., and Gasser S.M. 2004. Separation of silencing from perinuclear anchoring functions in yeast Ku80 Sir4 and Esc1 proteins. *EMBO J.* **23:** 1301–1312.

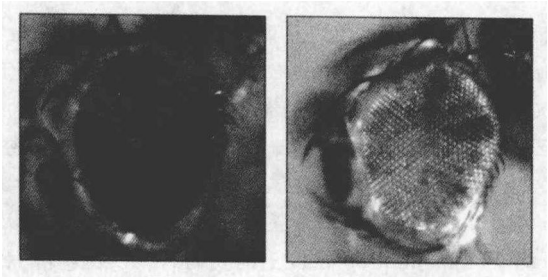
Tanner K.G., Landry J., Sternglanz R., and Denu J.M. 2000. Silent information regulator 2 family of NAD-dependent histone/protein deacetylases generates a unique product, 1-*O*-acetyl-ADP-ribose. *Proc. Natl Acad. Sci.* **97:** 14178–14182.

Thompson J.S., Johnson L.M., and Grunstein M. 1994. Specific repression of the yeast silent mating locus HMR by an adjacent telomere. *Mol. Cell. Biol.* **14:** 446–455.

van Leeuwen F., Gafken P.R., and Gottschling D.E. 2002. Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* **109:** 745–756.

Weiss K. and Simpson R.T. 1998. High-resolution structural analysis of chromatin at specific loci: *S. cerevisiae* silent mating type locus HMLα. *Mol. Cell. Biol.* **18:** 5392–5403.

# 第 5 章



# 果蝇位置效应花斑、异染色质形成和基因沉默

Sarah C. R. Elgin[1] and Gunter Reuter[2]

[1] *Department of Biology, Washington University, St. Louis, Missouri 63130*

[2] *Institute of Genetics, Biologicum, Martin Luther University Halle, D-06120 Halle, Germany*

## 本章目录

1. 基因异常移位到异染色质附近引起花斑表型
2. 对 PEV 抑制子和增强子的筛选得到了染色体蛋白和染色体蛋白修饰子
3. 多线染色体的免疫荧光定位鉴定出与异染色质特异结合的蛋白质
4. 组蛋白修饰在异染色质沉默中起到关键作用
5. 染色质蛋白形成相互依赖的复合体来维持和延伸异染色质结构
6. 果蝇异染色质是如何定位形成的
7. 异染色质并不都是一样的
8. 不同生物中的 PEV、异染色质形成和基因沉默
9. 总结：关于异染色质我们知之甚少

致谢

参考文献

## 概要

基因通过染色质重排或者转座作用移位到异染色质附近后呈现出花斑表型，表明正常情况下活跃表达的基因在部分细胞中被沉默了，这就是位置效应花斑（position-effect variegation，PEV）。位置效应花斑中的基因沉默是由于报告基因凝聚成异染色质而引起的，这表明异染色质一旦开始形成就能够蔓延到周围区域。在黑腹果蝇中可以进行遗传、细胞和生化等分析。在本章节中，我们将会介绍如何综合运用这些方法来发现位置效应花斑中的作用蛋白，并发现在异染色质的建立和维持过程中起关键作用的蛋白。异染色质的形成关键是组蛋白 H3 第 9 位赖氨酸（H3K9）的甲基化、异染色质蛋白 1（heterochromatin protein 1，HP1）的结合以及包括 H3K9 甲基转移酶在内的 HP1 相互作用蛋白的结合，这些蛋白之间的相互作用对于异染色质的维持和延伸是必需的。异染色质的形成的靶向（包括 H3K9 甲基化的积累）可能涉及 RNA 干扰（RNA interference，RNAi）的调控“机器”；特异的蛋白与 DNA 相互作用也可能起到一定的作用。尽管各类异染色质区域（着丝粒区、端粒和第四染色体）有着共同的生化特征，但是它们彼此之间还是不同的，着丝粒区的异染色质是镶嵌分布的。果蝇的异染色质区少有基因存在但并非没有基因，而且这些异染色质区基因的充分表达是依赖于这种异染色质环境的。一个诠释异染色质形成和维持（包括定位和蔓延）机制的最终模型，必须考虑到不同基因对这种染色质环境的反应。

## 1. 基因异常移位到异染色质附近引起花斑表型

一些真核生物基因组中大段大段 DNA（主要是重复序列）被压缩成永久失活的组成型异染色质。这种染色质组分的最初发现是因为它们从细胞分裂中期到分裂间期的转变过程中始终保持着浓缩和深染的状态（异固缩的）。它们主要分布在染色体的着丝粒区和端粒区。异染色质区域倾向于晚复制，在减数分裂期少发生或不发生染色体重组。这些区域虽缺乏基因，但并不是没有基因，且这些异染色质区基因的有效表达常常依赖于异染色质环境。约 1/3 的果蝇基因组被认为是异染色质，包括整个 Y 染色体、微小的第四号染色体的大部分以及占 X 染色体 40%和常染色体 20%的着丝粒区。在过去的几十年中，我们获得了大量关于异染色质的生物化学知识，而这其中的大部分是从果蝇的研究中得到的（Richards and Elgin 2002；Schotta et al. 2003）。

*white* 基因是在黑腹果蝇中最早发现的突变之一，它导致了果蝇的白眼而非野生型的红眼。Muller（1930）用 X 射线诱变发现了一种不寻常的果蝇表型，这些果蝇眼睛的局部区域是红色的而另外的区域是白色的（图 5-1）。这种表型意味着 *white* 基因本身并没有被彻底破坏，因为一部分眼睛还保持着红色，如再次用 X 射线诱变依然能够将其后代眼睛的颜色完全恢复成红色。然而，*white* 基因在有些细胞里显然是沉默的，在正常情况下 *white* 基因在这些细胞里是表达的。对多线染色体的检测（图 5-4）表明这种表型是染色体倒位或重排的结果，其中一个断裂点发生在臂间异染色质内而另一个断裂点则邻近 *white* 基因（图 5-1）。由于花斑表型是由基因的染色体上位置变化引起的，所以这种现象被称为位置效应花斑（PEV）。实际上，到目前为止被检测过的所有果蝇基因在染色体重排后都表现出异质性，凡是涉及臂间异染色质的染色体重排都能导致 PEV 的发生。PEV 的现象存在于包括酵母、果蝇和哺乳动物在内的各个物种中，但主要在果蝇中被用作研究异染色质的工具。

PEV 现象表明染色体重排导致异染色质构象沿着染色体“蔓延”。显然，这种染色体重排删除了正常存在的壁垒或缓冲区域。其后果就是正常情况下以常染色质形式存在的基因发生染色质结构改变被沉默。对带有这种染色体重排的果蝇幼虫多线染色体的视觉观察表明带有报告基因的区域被包装成一段致密的异染色质区域，但是这种现象只在报告基因失活的细胞中才存在（Zhimulev et al. 1986）。*white* 基因重排的表型模式，如色素细胞的数目、色素斑的大小和在两个不同细胞类型观察到的色素水平可以是不同的（图 5-1）。在一个以可诱导的 *lacZ* 基因作为报告基因的系统中，研究者们观察到基因沉默发生在胚胎期，这是一个在细胞学水平上可以观察到异染色质和异染色质在体细胞和生殖细胞中被表观遗传继承的最早时期；在三龄幼虫中，这种镶嵌表型是由于分化过程中基因沉默被不同程度地降低所造成的（Lu et al. 1996）。然而，不是所有异质的基因沉默都能保持到分化后，对不同的基因而言，决定其“开/关”平衡的因素也是不同的（更多的详细讨论见 Ashburner et al. 2005b）。

对于一只表现出 PEV 表型的果蝇来说，可以直接筛选出显性的新突变位点（可以通过化学诱变剂诱发点突变或小片段插入或删除），这些突变有的是导致基因沉默丢失

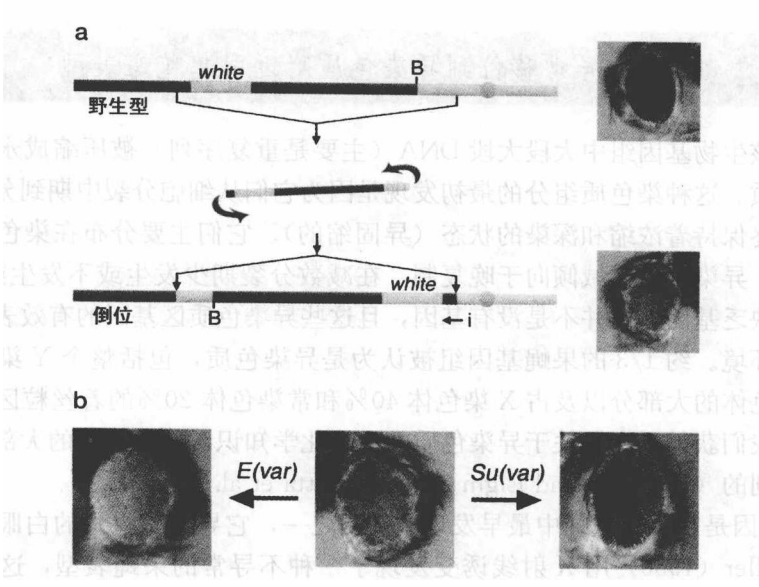



图 5-1 $In(1)w^{m4}$ 中 X 染色体倒置造成 *white* 基因花斑的示意图

*White* 基因座通常位于常染色体远端（蓝色），由于 X 射线诱导的染色体倒位发生于 X 染色体臂间异染色质 25kb 的断裂点内（粉红）。异染色质包装延伸到常染色质区域导致了基因沉默；有些细胞在分化过程中丢失了基因沉默就导致了花斑表型。对于表现出 PEV 现象的果蝇，人们发现了一些新突变，如抑制花斑表型的 *Su*（*var*）突变，导致基因沉默丢失；或增强花斑表型的 *E*（*var*）突变，从而引起基因沉默增强。

的 PEV 抑制子，称为花斑抑制子［*Su*（*var*）］，有的是导致基因沉默增强的 PEV 增强子，称为花斑增强子［*E*（*var*）］（图 5-1）。已有大约 30 个 PEV 修饰因子被分离和鉴定出来，估计还将有更多的候选蛋白出现。只要一个基因被克隆、其基因产物被鉴定，基本上就可以说是发现一个染色体蛋白或一个染色体蛋白的修饰因子（见后文所述）。这些基因位点中一部分既可以导致单倍体异常表型又可以导致相反的三倍体异常表型，例如，这个基因的单拷贝可以抑制 PEV，三拷贝则可增强 PEV。这些位点提示，这些基因的蛋白质产物在异染色质的结构中起作用，而且异染色质的蔓延受到这些蛋白质的剂量调控（图 5-2）（Locke et al. 1988）。然而，蔓延是一个复杂的过程，不是一个简单的线性延续，这可能主要取决于被沉默区域的 DNA 的组织情况。

染色体重排的结果暗示常染色质区的基因通过转座插入到异染色质区同样也会表现出 PEV，事实也正是如此。一种在许多野生果蝇种属中都发现的 DNA 转座子——*P* 元件（*P* element），可以用于这一研究目的。自然的 *P* 元件在每个末端都有特征性的反向重复序列，编码一种 *P* 元件特异的转座酶。含有目的基因但不含 DNA 转座酶的报告载体与编码有活性的转座酶载体通过混合注射的方式导入果蝇胚胎中后，也可以使报告载体整合到果蝇基因组中。一种以 *P* 元件为基础的转座元件携带一个拷贝的 *hsp70* 驱动的 *white* 基因，可被用于在不含内源拷贝 *white* 基因的果蝇中鉴定异染色质区域（图 5-3a）。当 *P* 元件被插入到常染色质区中，果蝇的眼睛是红色的；当这个 *P* 元件移位后（通过引入编码转座酶基因），有近 1%的果蝇表现出 PEV。原位杂交实验表明这些例子

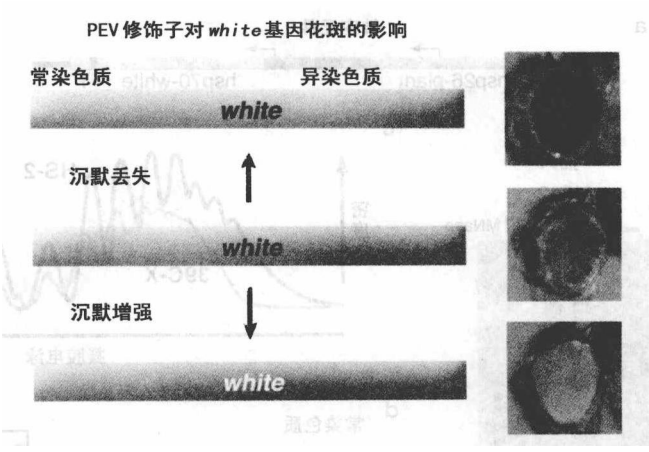



图 5-2　一些 PEV 修饰子的剂量依赖效应

具有剂量依赖效应的 PEV 修饰子被认为是异染色质结构蛋白。野生型修饰子基因以双拷贝形式存在时（中图）的花斑表型（这里表示为 *white* 报告基因），在修饰子基因以三拷贝形式出现的情况下，可促使更广泛的异染色质形成、增强了报告基因的沉默（下图）。相反，当野生型修饰子基因以单拷贝形式存在时，将会导致更少的异染色质形成以及报告基因的更多的表达（上图）。

中的 *P* 元件“跳跃”到了臂间异染色质、端粒或者小的第四号染色体中（Wallrath and Elgin 1995）。这种异染色质区域的鉴定与早期的细胞学研究是一致的。

这种 *P* 元件的应用允许人们比较同一个报告基因在常染色质环境和异染色质环境中不同的装配状态。不管是非特异性的（如 Dnase I）还是特异性的（限制性内切核酸酶），异染色质相对更能抵抗核酸酶的切割，并且外源探针也更不容易接近异染色质，如 *dam* 甲基转移酶区域。用微球菌核酸酶（MNase）对插入在常染色质和异染色质环境中相同的 *hsp26* 转基因（以单一的植物 DNA 片段标注，图 5-3a）进行分析，结果表明在异染色质中的转基因的核小体排列变得更加有秩序了，意味着异染色质中的核小体排列有着更加规则的空间间隔（图 5-3b，c）。MNase 切割的片段的界线是清晰的，表明 MNase 切割位点比通常在连接区中的更小了。核小体的规则排列一直蔓延到基因的 5′调控区，这无疑有助于解释观察到的 5′超敏感位点缺失的现象（Sun et al. 2001）。的确，尽管基因沉默的机制还没有完全被研究清楚，但是已经有了大量关于高度花斑基因的转录抑制的证据，包括 TFIID 与其他转录因子结合的丢失（Cryderman et al. 1999b）

## 2. 对 PEV 抑制子和增强子的筛选得到了染色体蛋白和染色体蛋白修饰子

PEV 能够被许多因子修饰。发育期间的温度和基因组内异染色质的总量被证明是影响花斑程度的首要因素。一般来讲，发育期间的温度升高（25℃～29℃）抑制花斑效应，而发育期间温度的降低（如 18℃）则增强花斑效应（增强基因沉默）。其他能够加速或者延缓发育的因素也能起到类似的作用。在携带有一条多余 Y 染色体的果蝇（XXY 雌和 XYY 雄）中，PEV 被明显抑制，而在没有 Y 染色体的雄果蝇（XO）中，PEV 明显增强。总的来说，双倍的异染色质含量抑制花斑效应，而去除异染色质则增

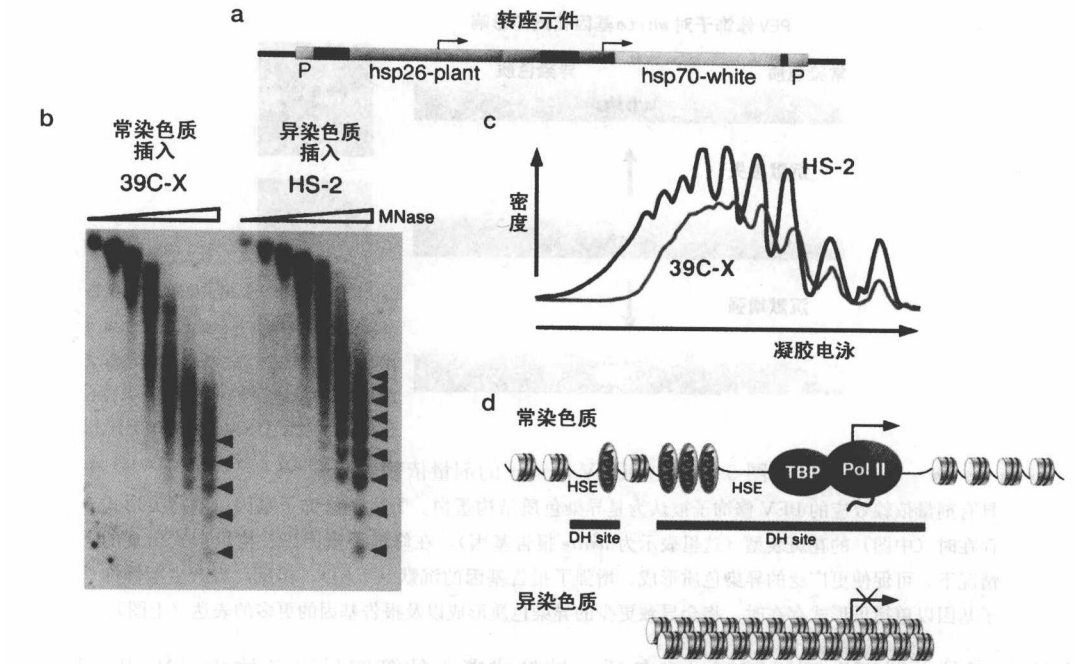



图 5-3 异染色质被装配成规则的核小体阵列

a. 一个转座元件，携带一个拷贝的热激蛋白（hsp26）基因和一个拷贝的 hsp70 驱动的 *white* 基因作为可视标记，用于检查处于不同染色质区域的相同基因。b. 携带处于常染色质区转基因的果蝇系（39C-X；红眼）和携带处于异染色质区转基因的果蝇系（HS-2；花斑眼）果蝇胚胎的核被浓度逐渐增加的 MNase 切割，纯化 DNA 进行琼脂糖凝胶电泳，然后用针对转基因的探针进行 Southern 杂交。MNase 切割的连接区位点用箭头标记。c. 对每个样品的最后一个泳道用光密度计扫描结果的对比（从上到下即是从左到右）。在异染色质中能检测到 9～10 个核小体的阵列（红线），而在常染色质中是 5～6 个核小体的阵列（蓝线），提示在异染色质中有更规则的核小体阵列的空间间隔。d. 结构示意图（b，c，经允许修改自 Sun et al. 2001 [©美国微生物学协会]）。

强花斑效应。这种影响也许是因为具有固定数量的异染色质装配关键蛋白的滴度发生变化造成的。抑制或增强 PEV 的第一批位点突变体是由 Schultz（1950）和 Spofford（1967）鉴定出来的。截至目前，约有 150 个基因被认为是 PEV 修饰子。

*Su(var)*和 *E(var)*突变筛选出来的基因与 PEV 中异染色质基因沉默的起始有关。对这些基因的分子生物学研究是理解异染色质形成和基因沉默机制的关键。在大多数情况下，这些突变对 PEV 的修饰效应是显性的，*Su(var)*/＋或 *E(var)*/＋杂合子表现出 PEV 抑制或增强表型（图 5-1）。*Su(var)*和 *E(var)*突变的有效分离和深入的遗传学分析取决于是否有可用于实验操作的 PEV 重排。尽管已有大量的 PEV 重排报道（Flybase 2005），仅有少部分可被有效地用于显性修饰子突变的遗传学筛选。对这类研究最有用的 PEV 重组是 $In(1)w^{m4}$（Muller 1930）。这个重排是 *white* 基因花斑，是一个在成虫果蝇中容易识别的表型（图 5-1）。在 $w^{m4}$ 中 *white* 基因花斑的外显率是 100%，因此所有原始的果蝇都表现出 *white* 基因花斑表型的眼睛。*white* 基因的失活不影响果蝇的生活力和生育力，所以用 $w^{m4}$ 纯合子果蝇可开展各种相关研究工作。

在 $w^{m4}$ 重排中，染色体倒位造成携带 X 染色体异染色质的 *white* 基因位于核仁组织的远端边界区（Cooper 1959）。这个区域含有串连排列的 R1 型移动元件；*In(1)* $w^{m4}$ 异染色质区的断裂点正好位于一个 R1 重复单元中（A. Ebert and G. Reuter，未发表）。经 X 射线或甲基磺酸乙酯（ethyl methylsulfonate，EMS，一种化学诱变剂）处理后已分离到 *In(1)* $w^{m4}$ 的 $W^+$ 表型回复突变体（Tartof et al. 1984；Reuter et al. 1985）。在对超过 50 个 $W^+$ 表型回复突变体的染色体（全都是再导致 *white* 基因转座到常染色质附近区域）的分析表明，断裂点周围的异染色质导致了 $w^{m4}$ 的 *white* 基因的失活。如果引入强 *E(var)*突变，大部分的回复突变体再次表现出花斑表型，意味着重定位后部分异染色质序列仍然与 *white* 基因联系在一起。这并不奇怪，因为周围 DNA 中的断裂点是随机的。这项研究意味着重复序列（本项研究是 R1 重复单元）是异染色质形成的靶点。

大部分已知的 PEV 修饰子突变是在敏感的遗传背景下被分离出来的。在显性抑制子突变的分离中，检测到的基因型中含有显性增强子，而显性抑制子则往往出现于显性增强子突变的分离中（Dorn et al. 1993b）。如果检测基因型中含有 PEV 的增强子，所有的 $w^{m4}$ 果蝇都应为白眼，如果只有个别的带有花斑或者红眼，表明有新的显性 *Su(var)*突变介入。相应地，在含有显性抑制子的测试果蝇中，所有的 $w^{m4}$ 果蝇都应为红眼，如果只有个别的为花眼，表明有新的显性 *E(var)*突变介入。这些敏感的遗传学筛选方法有助于分离强显性 *Su(var)*和 *E(var)*突变，对详细的遗传学分析是非常有用的。

至少有 100 万只果蝇被应用于采用这种方法的各种筛选中，已有超过 140 个 *Su(var)*突变和230个 *E(var)*突变被分离了出来（Schotta et al. 2003）。突变是通过 EMS，X 射线处理或 *P* 元件重定位实现的。另一组 *Su(var)*突变是在对 $w^{m4}$ 果蝇的直接筛选中分离出来的（Sinclair et al. 1983）。在对体色标记基因 *yellow* 高度花斑效应并含有 Df（1；f）染色体的果蝇进行筛选之后，也分离出了 70 个 PEV 修饰子突变（Donaldson et al. 1994）。另外，对转座子报告基因表达的显性修饰子的筛选也鉴定出一些带有 *Su*（*var*）效应的突变（Birchler et al. 1994）。有相当一批关键调控基因的表达被 polycomb 家族（PcG）下调，或被 trithorax 家族（trxG）上调。在直接检测中，只有很少的 *PcG* 突变影响 PEV（如 Sinclair et al. 1998）。而很多 *trxG* 基因的突变是 PEV 的增强子（Dorn et al. 1993a；Farkas et al. 1994）。

总的来说，这些筛选鉴定出了大量的显性 *Su(var)*和 *E(var)*突变。据当前的遗传学分析来看，*Su(var)*和 *E(var)*突变的总数大约是 150 个。大量的 *Su(var)*和 *E(var)*基因有着几乎完全相同的表型，这有时会造成遗传学命名的矛盾。所以通常情况下，*Su(var)*和 *E(var)*基因符号是突变所在染色体号的数字、基因的排序数和等位基因的排序数的综合。例如，*Su(var)3-9*$^{17}$表示在 3 号染色体上发现的第 9 个 *Su(var)*基因的第 17 个等位基因。目前，只有 30 个左右的相应基因被精确定位且等位基因都被鉴定出来（表 5-1）。利用一系列重叠缺陷以及复制加倍观察到大约有三分之一已鉴定的基因有剂量依赖效应。在这些例子中，基因丢失一个拷贝造成了基因产物量的减少，进而影响了花斑表型。这些 *Su(var)*和 *E(var)*位点的删除分别造成基因沉默的抑制或增强。染色体重复片段研究鉴定出了一些修饰子位点，如果通过复制或转基因插入的方法引入这些

基因的一个额外拷贝则会表现出对 PEV 完全相反的修饰作用。具有剂量依赖效应的 PEV 修饰子基因的总数估计为 15～20 个（Schotta et al. 2003）。

**表 5-1 遗传学上定义的 *Su(var)* 和 *E(var)* 基因以及它们的分子功能**

| *Su(var)*/ *E(var)*基因 | 染色体位置 | 分子功能、蛋白定位和表型作用 |
|---|---|---|
| *Suv4-20*[*Su(var)*] | X; 1B13-14 | HKMT，三甲基化组蛋白 H4K20 |
| *Su(z)5*[*Su(var)*] | 2L; 21B2 | S-腺苷硫氨酸合成酶 |
| *chm*(*chameau*)[*Su(var)*] | 2L; 27F3-4 | Myst 结构域 HAT；抑制 PEV 但是增强 Poly-group 突变 |
| *Su(var)2-5*(HP1) | 2L; 28F2-3 | 异染色质蛋白 HP1，结合二甲基化和三甲基化的 H3K9，结合 SU(VAR)3-9 |
| *Su(var)2-HP2* | 2R; 51B6 | 异染色质结合蛋白，结合 HP1 |
| *Su(var)2-10* | 2R; 45A8-9 | PIAS 蛋白，JAK/STAT 通路的负调节子 |
| *Su(var)3-64B*(HDAC1=RPD3) | 3L; 64B12 | 组蛋白去乙酰化酶 HDAC1，去乙酰化 H3K9 |
| *E(z)*[*Su(var)*] | 3L; 67E5 | HKMT，H3K27 单、双和三甲基化；增加一个基因拷贝则增强 PEV；在无效突变体中，所有常染色质和异染色质的 H3K27 甲基化都丢失了，H3K9 甲基化不受影响 |
| *SuUR*[*Su(var)*] | 3L; 68A4 | 抑制异染色质复制不足；异染色质结合蛋白 |
| *Su(var)3-1*(*JIL1*) | 3L; 68A5-6 | 反效等位基因的 *JIL1* 突变，羧基端缺失蛋白不影响激酶活性；阻止异染色质延伸 |
| *Dom*(*Domina*)[*Su(var)*] | 3R; 86B1-2 | fork head winged-helix(FKH/WH)蛋白；异染色质结合 |
| *Su(var)3-6* | 3R; 87B9-10 | PP1 蛋白丝氨酸/苏氨酸磷酸酶 |
| *Su(var)3-7* | 3R; 87E3 | 锌指蛋白，结合异染色质；与 HP1 和 SU(VAR)3-9 相互作用 |
| *Su(var)3-9* | 3R; 89E6-8 | HKMT，甲基化组蛋白 H3K9，结合异染色质，与 HP1 相互作用 |
| *mod*(*modulo*)[*Su(var)*] | 3R; 100E3 | DNA 和 RNA 结合蛋白，磷酸化的 Mod 结合 rRNA |
| *E(var)3-64E*/ *Ubp64*$^{Evar1}$ | 3L; 64E5-6 | 可能的泛素特异的蛋白酶(Ubp46) |
| *Trl*(*Trithorax-like*)[*E(var)*] | 3L; 70F4 | GAGA 因子，结合重复 DNA 序列 |
| *Mod*(*mdg4*)/ *E(var)3-93D* | 3R; 93D7 | 转录调节因子，通过反式剪切产生 20 多个蛋白亚型 |
| *E(var)3-93E* | 3R; 93E9-F1 | E2F 转录因子；单倍增强子和三倍抑制子 |

如果一个基因丢失一个拷贝抑制了 PEV，而同时存在三个拷贝则增强了 PEV，这意味着这个基因编码的产物与异染色质的建立和基因的沉默是一种线性量化关系（图 5-2）。3 个这样的位点，*Su(var)2-5*（编码 HP1），*Su(var)3-7*（编码一个锌指蛋白）和 *Su(var)3-9*（编码一个组蛋白赖氨酸转移酶）都被鉴定出来。*Su(var)2-5* 是对一个 cDNA 表达文库进行筛选时克隆出来的，它的单克隆抗体识别异染色质（James and

Elgin 1986)。它编码的异染色质相关蛋白被命名为 HP1（异染色质蛋白 1)。用分离克隆的 DNA 做原位杂交分析鉴定出一个位于多线染色体 28～29 区的基因，这是 *Su(var) 2-5*以前被定位的地方。对突变等位基因的 DNA 序列分析证实位于染色体 28F1-2 的 *Su (var) 2-5*位点编码 HP1（Eissenberg et al. 1990)。HP1 包含两个保守结构域，一个 N 端 chromo 结构域和一个 C 端 chromoshadow 结构域（Paro and Hogness 1991)，并与其他几个染色质蛋白相互作用。*Su(var) 3-7*最先在细胞遗传学上被定位（用一系列重叠删除和重复）在第 3 号染色体的 87E1-4 区。这个区域曾在部分最初的果蝇染色体步移实验中于 DNA 水平被分析过（Bender et al. 1983)。用一系列重叠的基因组克隆，*Su(var) 3-7*被定位于一个 7.8kb 的 DNA 片段，对花斑报告基因有三倍增强子作用（Reuter et al. 1990)。*Su(var) 3-7*编码一个具有 7 个规则间隔的锌指结构的蛋白，这个结构域具有 DNA 结合功能（Cleard and Spierer 2001)。*Su(var) 3-9*在 *P* 元件转座子标签实验中被克隆出来（Tschiersch et al. 1994)。*Su(var) 3-9*在果蝇中（以及在截至目前研究过的其他全变态昆虫中）与编码 eIF2γ 的基因形成双顺反子单元（Krauss and Reuter 2000)。由于 *Su(var) 3-9*转录单元中没有内含子，所以 *Su(var) 3-9*很有可能是通过逆转座作用插入到 *eIF2γ* 基因的内含子中。SU(VAR)3-9蛋白在 N 端含有一个 chromo 结构域，在 C 端含有一个 SET 结构域，这是一个首先在 SU(VAR) 3-9、[E (Z)]和 trithorax 中鉴定出来的结构域（Jones and Gelbart 1993; Tschiersch et al. 1994)，是一个特异修饰组蛋白 H3 第 9 位赖氨酸的组蛋白甲基转移酶。

*Su(var) 2-5*已发现了 7 个不同的突变体等位基因，包括 chromo 结构域的错义突变，终止密码子提前和剪切错误（Eissenberg et al. 1992)。*Su(var) 3-7*的突变体是借助于同源重组产生的（Seum et al. 2002)；其他的等位基因是作为 *P* 元件依赖的基因沉默的抑制子被发现的（Bushey and Locke 2004)。*Su(var) 3-9*有 40 个突变体等位基因在分子水平上被发现（Ebert et al. 2004)。用特异抗体和转基因表达的融合蛋白进行的免疫组化分析表明这三个蛋白都倾向于与异染色质结合（图 5-4）（James et al. 1989; Cleard et al. 1997; Schotta et al. 2002)。HP1 和 SU（VAR）3-9 与异染色质的共定位尤其明显。这些蛋白同样也结合在端粒和一些常染色质位点上（Fanti et al. 1998; Schotta et al. 2002)。

数个位于端粒区携带有 $w^+$ 报告基因的 *P* 元件表现出 *white* 基因的花斑效应。这个现象被称为端粒位置效应（telomere position effect, TPE)。组成果蝇端粒的端粒相关卫星序列（telomere-associated satellite, TAS)、紧邻 HeT-A 和 TART 逆转录病毒元件的重复 DNA 簇呈现异染色质样装配（Cryderman et al. 1999a)。总的来说，HP1 对端粒的完整性很重要，但发生在已知修饰因子基因中的突变并不影响 TPE。在缺失 HP1 的细胞中，染色体常常发生端粒融合（Fanti et al. 1998)。在果蝇中进行的筛选中，没有鉴定出反式作用于 TPE 的显性修饰因子（Mason et al. 2004)，意味着这些区域是被两个或更多的独立机制沉默的。

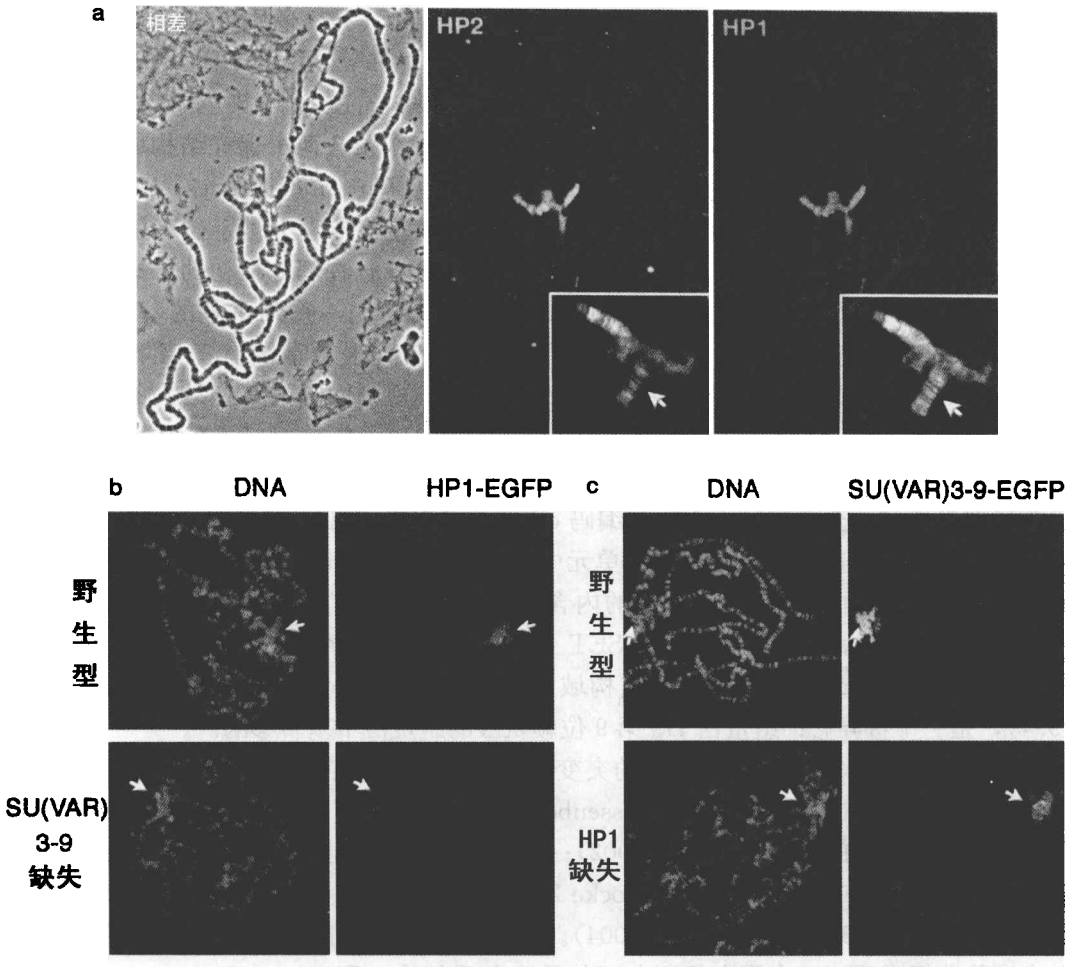



图 5-4 多线染色体的免疫荧光染色鉴定出主要与异染色质结合的蛋白

(a) 果蝇幼虫唾腺的多线染色体的相差显微镜拍照结果（左图），先与针对某一染色质蛋白的抗体孵育，再与偶连免疫荧光标签的二抗孵育。HP1（右图）和 HP2（中间图）有着相似的分布态势，主要分布在臂间异染色质，短小的第 4 号染色体（内图，箭头所指）以及一部分常染色质臂位点。值得注意的是，抗体的效率可能会受到固定方法的影响（见 Stephens et al. 2003）。(b, c) HP1 和 SU(VAR)3-9与臂间异染色质的结合是相互依赖的。Su(var)3-9的突变导致与臂间异染色质结合的 HP1 的丢失（但不是第 4 号染色体，见正文）(b)，而 Su(var)2-5的突变则导致 SU(VAR)3-9的脱离 (c)（修改自 Shaffer et al. 2002）。

## 3. 多线染色体的免疫荧光定位鉴定出与异染色质特异结合的蛋白质

用果蝇中做研究的一个优点是可以检测多线染色体，提供可视的基因组全景。在幼虫期，许多终端分化细胞的染色体虽进行复制但不进行有丝分裂；且染色单体在联会中所有拷贝保持配对。最极端的例子发生在唾腺中，染色体的常染色质臂经历 10 轮复制，产生大约 1000 个拷贝。然而复制不是完全一样的，很多重复序列保持低水平复制，卫星 DNA 序列根本就不复制。所有的染色体臂融合在共有的染色中心。因此，在黑腹果

蝇中，人们可以观察到 5 个长臂——X 染色体，第 2 号染色体左臂［2L］，第 2 号染色体右臂［2R］，第 3 号染色体左臂［3L］，第 3 号染色体右臂［3R］，以及由臂间异染色质组成的浓缩的染色中心延伸出来的一条短小的第 4 号染色体（图 5-4a）（综述见 Ashburner et al. 2005）。

尽管遗传分析鉴定出许多异染色质形成所需要的基因位点，但是它本身并无法确定基因位点的产物对异染色质形成起直接或是间接的作用。与异染色质特异结合的蛋白是通过筛选单克隆抗体（由一部分紧密结合的核蛋白产生的）在多线染色体上的分布得到的。在免疫荧光染色实验中，22kD 蛋白（随后被命名为 HP1）的特异抗体分布在臂间异染色质、端粒和小的第 4 号染色体带的部分，这都是异染色质区（图 5-4a）（James and Elgin 1986）。接下来的研究（如上所述）证明 HP1 是由 *Su(var)2-5* 编码的，是 PEV 的抑制子（Eissenberg et al. 1990）。用特异抗体检查蛋白在有丝分裂的染色体或者多线染色体上的分布仍然是对 *Su(var)* 编码的蛋白是否为染色体蛋白的最好证明。有近 10 个这样的异染色质特异的蛋白被鉴定出来；如果编码这些蛋白的基因发生突变，人们常常会观察到对 PEV 的显性抑制（表 5-1）（Ashburner et al. 2005b）。这些蛋白，包括最近发现的 HP2（图 5-4a）（Shaffer et al. 2002），都很有可能是异染色质的结构组分。

## 4. 组蛋白修饰在异染色质沉默中起到关键作用

对 SU(VAR)3-9的分析发现它对异染色质基因沉默起关键作用（Tschiersch et al. 1994）。SU(VAR)3-9含有一个在组蛋白 H3K9 甲基化中起催化作用的 SET 结构域。SU(VAR)3-9是 H3K9 特异的组蛋白赖氨酸甲基转移酶。这个结论是在对 SU(VAR)3-9 的人同源蛋白 SUV39H1 进行的鉴定中率先做出的（Rea et al. 2000）。在果蝇中，SU(VAR)3-9是主要 H3K9 组蛋白甲基转移酶（histone methyl-transferase，HKMT）但并不是唯一的（Schotta et al. 2002；Ebert et al. 2004）。SU(VAR)3-9主要调控臂间异染色质的二甲基化，但不调控第 4 号染色体、端粒和常染色质位点的甲基化。在果蝇中主要位于染色中心内部的 H3K9 三甲基化也是由 SU(VAR)3-9控制的。这个区域的二甲基化以及臂间异染色质的单甲基化却不是由 SU(VAR)3-9控制的（Ebert et al. 2004）。负责这些修饰的 HKMT 还没有被发现。SU(VAR)3-9对 PEV 强烈的剂量依赖作用证明了 H3K9 二甲基化在异染色质基因沉默中的重要性（见上面的讨论），SU(VAR)3-9的突变与基因沉默的抑制和它的 HKMT 活性有关也证明了这一点。酶学上，超活化的 Su(var)3-9$^{ptn}$突变体是 PEV 的强抑制子，造成染色体中心 H3K9me2 和 H3K9me3 水平的提高，在很多常染色质位点上也产生了明显的 H3K9me2 信号（异位异染色质）。S-腺苷甲硫氨酸是所有这些甲基化反应中甲基的供体，编码 S-腺苷甲硫氨酸合成酶的基因的突变体 *Su(z)5* 也是 PEV 的显性抑制子（Larsson et al. 1996）。

利用 *Su(var)* 基因突变的研究已开始逐步揭示异染色质建立所需分子反应的顺序了。SU(VAR)3-9结合在异染色质序列既依赖其 chromo 结构域也依赖其 SET 结构域（Schotta et al. 2002）。SU(VAR)3-9的结合是如何被控制的机制尚未清楚。SU(VAR)

3-9介导的 H3K9 甲基化建立了 HP1 的结合位点。HP1 的 chromo 结构域可特异地结合 H3K9me2 和 H3K9me3（Bannister et al. 2001；Lachner et al. 2001）。SU(VAR)3-9与 HP1 结合已被酵母双杂交实验和免疫共沉淀实验所证实（Schotta et al. 2002）。SU(VAR)3-9的 N 端和 chromo 结构域与 HP1 的 chromoshadow 结构域相互作用。SU(VAR)3-9的这个区域同样与 SU(VAR)3-7的羧基端结构域相互作用。SU(VAR)3-7 在 3 个不同的位点与 HP1 的 chromoshadow 结构域相互作用（Delattre et al. 2000）。由于这些相互作用，人们可以推测这三个蛋白——HP1、SU(VAR)3-7和 SU(VAR)3-9是在由多个蛋白组成的异染色质蛋白复合体中发生相互作用。

HP1 和 SU(VAR)3-9与臂间异染色质的结合是相互依赖的（Schotta et al. 2002）。SU(VAR)3-9引起 H3K9 的二甲基化，二甲基化的 H3K9 又特异地被 HP1 的 chromo 结构域识别（Bannister et al. 2001；Lachner et al. 2001）。因此，在 *Su(var)3-9* 缺失的幼虫中，HP1 与臂间异染色质的结合削弱了（图 5-4b）。这反映了 HP1 结合的特异性，它结合 H3K9me2 但却不结合 H3K9me1；单甲基化不受 SU(VAR)3-9影响（Ebert et al. 2004）。染色体中心内部区域、第 4 号染色体、端粒以及一些常染色质位点 H3K9 的二甲基化并不依赖于 SU(VAR)3-9，所以在 SU(VAR)3-9突变系中，HP1 依然结合在这些位点上。在野生型细胞中，SU(VAR)3-9结合在这些位点上，但是失活的；一种未知的 HKMT 控制着这些位点的 H3K9 甲基化。

相反，如果 HP1 不存在（突变失活），SU(VAR)3-9将不再与臂间异染色质结合，而依然结合在染色体的常染色质臂上（图 5-4b）。在这种情况下，SU(VAR)3-9几乎结合到所有的带上，导致 H3K9 的单甲基化和二甲基化（H3K9me1 和 H3K9me2）（图 5-5）。因此，HP1 对于限制 SU(VAR)3-9在臂间异染色质上的结合是至关重要的。这些数据提示了这一系列反应是 SU(VAR)3-9先结合在异染色质区域上然后造成 H3K9me2。这一标记可被 HP1 的 chromo 结构域识别；SU(VAR)3-9与 HP1chromoshadow 结构域的相互作用保证了 HP1 与异染色质的结合。嵌合体蛋白 HP1-PC 是由 polycomb 蛋白（PC）的 chromo 结构域替代 HP1 的 chromo 结构域而产生的（Platero et al. 1996）。PC 的 chromo 结构域紧密结合在 H3K27me3 上（Fischle et al. 2003）。因此，嵌合体蛋白 HP1-PC 识别常染色质臂上的 polycomb 结合的 H3K27me3 位点；在有这样一个嵌合体蛋白 HP1-PC 存在的情况下，SU(VAR)3-9也存在于 PC 结合的位点上，证明 SU(VAR)3-9与 HP1 的 chromoshadow 结构域有紧密的相互作用（Schotta et al. 2002）。

在 SU(VAR)3-9缺失的细胞中，另一个异染色质特异的甲基化标记——H4K20 三甲基化（H4K20me3）也显著减少了（Schotta et al. 2004）。H3K9 二甲基化和 H4K20 三甲基化在异染色质中的相互依赖反应了 SU(VAR)3-9、HP1 和 SUV4-20 的相互作用。SUV4-20 是一个控制异染色质中 H4K20 甲基化的 HKMT。这个异染色质特异的甲基化标记在 SU(VAR)3-9缺失的细胞和 HP1 缺失的细胞中都被显著地削弱了，意味着在一个相互依赖的蛋白复合体中 SU(VAR)3-9、HP1 和 SUV4-20 是相联系的。*Suv4-20* 基因的突变引起 PEV 诱导的基因沉默的强烈抑制，意味着 H4K20me3 是这个过程所需要的。

第三个与异染色质形成有功能型联系的组蛋白甲基化标记是 E(Z)催化的 H3K27 甲

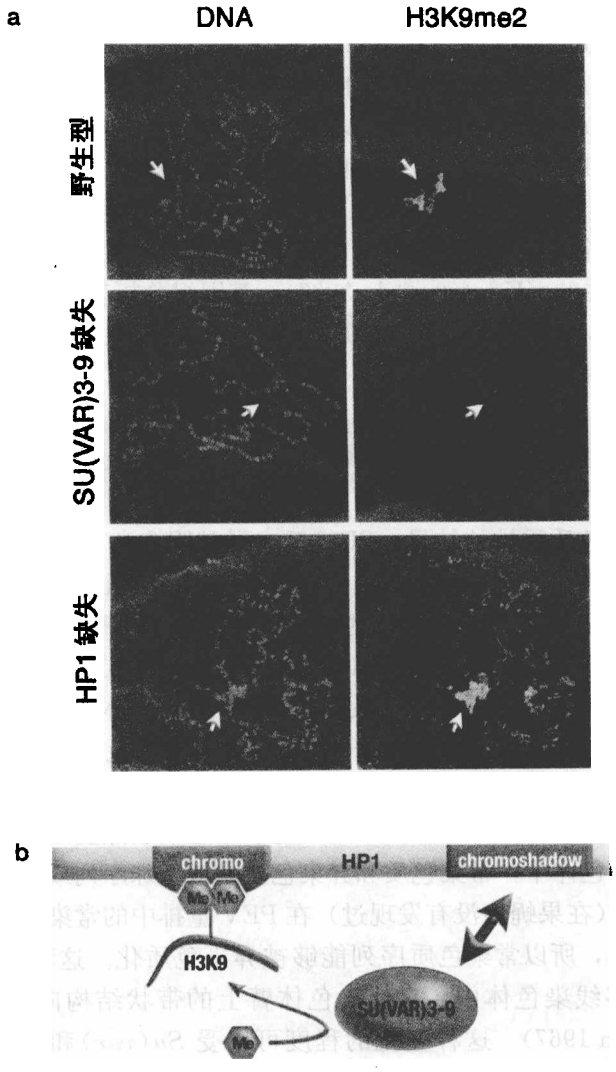



图 5-5　SU(VAR)3-9和 HP1 在决定 H3K9me 分布态式中的相互作用

(a) SU(VAR)3-9负责 H3K9 的二甲基化（H3K9me2）；缺失这个酶导致臂间异染色质该修饰的缺失，图中显示了多线染色体中抗体染色的缺失（比较中间一排图和上面一排图）。HP1 的缺失导致 SU(VAR)3-9定位的缺失；在整个染色体臂上都可见高水平的 H3K9me2（下排）。(b) HP1 通过 chromo 结构域与 H3K9me2 相互作用，通过 chromoshadow 与 SU(VAR)3-9相互作用。通过识别组蛋白修饰和负责这个修饰的酶，HP1 提供了一个异染色质延伸和表观遗传上遗传的机制。

基化。在果蝇中，E(Z)控制着常染色质中和异染色质中所有的 H3K27 的单甲基化、二甲基化和三甲基化。因此，在 E(Z)缺失的细胞中，H3K27 的甲基化消失了（Ebert et al. 2004）。E(Z)功能缺失突变的 Su(var)作用和额外的 E(Z)拷贝所带来的 PEV 增强子作用均暗示了 H3K27 在异染色质基因沉默中的功能（Laible et al. 1997）。这种作用是直接还是间接的尚未清楚。H3K27 甲基化在调控常染色质区 polycomb 沉默系统中起到关键作用。PC 和 HP1 在分布和功能方面少有重叠。至于 H3K27 甲基化如何参与 HP1

依赖的异染色质复合体，还有待于进一步阐述。在异染色质形成和相关的基因沉默、与H3K9me2和H3K9me3的结合以及与SU(VAR)3-9（H3K9的HKMT）、SUV4-20（H3K27的HKMT）和其他几个蛋白的相互作用中，HP1起到中心连接的功能。考虑到鉴定出来的Su(var)位点已经很多，可以肯定这个模型会越来越复杂。在动物和植物中，组蛋白H3K9和DNA甲基化是代表抑制状态染色质相互联系的标记（Martienssen and Colot 2001；Bird 2002）。多年以来，在果蝇基因组中是否有甲基化发生或根本就没有甲基化发生一直是个争论的焦点。最近的一些研究报告表明在早期的胚胎中是有低水平的DNA甲基化存在的（Kunert et al. 2003）。基因组分析表明，唯一可识别的DNA甲基转移酶是Dnmt2，但这个基因的突变对生物体的影响很小，不过，不能排除其在早期胚胎发育中的作用。

## 5. 染色质蛋白形成相互依赖的复合体来维持和延伸异染色质结构

作为遗传学上染色质重排的结果，PEV反应了基因表达的变化，特别是在一些细胞中报告基因表达丢失，而正常情况下这些基因是表达的。几个不同的模型被用来解释PEV，这些模型并不全部都是相互排斥的。最初被考虑的一个可能性是基因的随机丢失，这可能是晚复制造成的（Karpen and Spradling 1990）。定量Southern印迹分析表明这种解释是不可行的，表现出花斑效应的基因在二倍体组织中得到了充分复制（Wallrath et al. 1996）。其他的模型聚焦于花斑基因与异染色质区域的联系和（或）异染色质结构从新邻近的异染色质开始的蔓延。蔓延模型是建立在广泛的遗传学和细胞学数据上的，它解释了基因沉默是由于异染色质包装通过断裂点蔓延到常染色质的区域所造成的。在正常的染色体中，常染色质和异染色质是被特殊序列或缓冲区分隔开的。由于这些“绝缘序列”（在果蝇中没有发现过）在PEV重排中的常染色质与异染色质连接区并不存在（图5-1），所以常染色质序列能够被异染色质化。这种异染色质化在细胞学上是可见的，在多线染色体中，它使染色体臂上的带状结构向无定形的结构转变（Hartmann-Goldstein 1967）。这种变化的程度可以受*Su(var)*和*E(var)*突变的影响（Reuter et al. 1982）。

沿染色体臂长距离地抑制常染色质基因可以在遗传学上得到证明（Demerec and Slizymska 1937）。受影响的区域与HP1相结合，并且显示出H3K9me2的一个典型的标记（Ebert et al. 2004）。由于延伸模型假设了在这个过程中异染色质包装和常染色质包装相互竞争的关系，所以PEV修饰因子基因编码蛋白的功能要么控制异染色质的形成，要么控制常染色质的形成。正如上面所讨论的，剂量依赖的修饰因子的恢复支持了这个模型（Locke et al. 1988；Henikoff 1996）。最近，控制常染色质和异染色质平衡的*Su(var)*突变得到鉴定（Ebert et al. 2004）。PEV重排给了我们一个观察和研究异染色质包装向周围常染色区蔓延的例子。蔓延作用显然依赖于常染色质区内的一系列分子反应。已知几种染色质修饰是相互排斥的，它们界定了这两种不同的染色质状态。H3K9乙酰化、H3K4甲基化和H3S10磷酸化是活跃的常染色质的典型标记，而H3K9、H3K27和K4K20的甲基化则是沉默区域的特异标记。因此，常染色质区域的异染色质

化需要常染色质内组蛋白特异的去乙酰化、去甲基化和去磷酸化（图 5-6）。这个转换过程一开始依赖于 HDAC1 对 H3K9 的去乙酰化。*rpd3* 基因编码组蛋白 H3K9 特异的去乙酰化酶 HDAC1。*rpd3* 突变是 PEV 的强抑制子（Mottus et al. 2000），可拮抗 SU(VAR)3-9在基因沉默中的作用（Czermin et al. 2001）。HDAC1 与 SU(VAR)3-9/HP1 复合体相结合；这两个酶协同甲基化了原本乙酰化的组蛋白。

最近发现，异染色质向常染色质的蔓延在 *Su(var)3-1* 突变体中被完全阻断了（Ebert et al. 2004）。*Su(var)3-1* 突变是 JIL-1 激酶基因发生阅读框移码突变，造成表达的 JIL-1 是残缺的，即缺少 C 端区域。JIL-1 蛋白包含两个激酶结构域，其催化常染色质中 H3S10 的磷酸化。JIL-1$^{Su(var)3-1}$突变并不影响 H3S10 的磷酸化，但是可能影响了 H3S10 的去磷酸化，从而有效地抑制了 H3K9 的甲基化。这意味着这个过程还涉及一个磷酸酶。PP1 曾在 *Su(var)3-5* 突变中被鉴定出来（Baksa et al. 1993），它是否直接参与了这个过程还不得而知。H3K4 的去甲基化看起来似乎是常染色质区异染色质化的另一个必需条件。最近的研究工作表明，LSD1 氨基氧化酶是哺乳动物中 H3K4 的去甲基化酶（Shi et al. 2005）。推测的果蝇中 LSD1 同源蛋白 SU(VAR)3-3在所有测试的 PEV 重排中都促进了异染色质向常染色质区蔓延（S. Lein et al.，未发表）。与此相一致，在 *SU(VAR)3-3* 缺失的细胞中，断裂点附近常染色质的 H3K9 甲基化都消失了，但组成型异染色质却没有受到影响。这些结果证明，常染色质向异染色质的转换需要移除这些常染色质特异的组蛋白修饰标记，而这又需要几种不同的酶协同作用（图 5-6）。这些酶会与 SU(VAR)3-9/HP1 形成复合体，正如已经证明的 HDAC1 会与

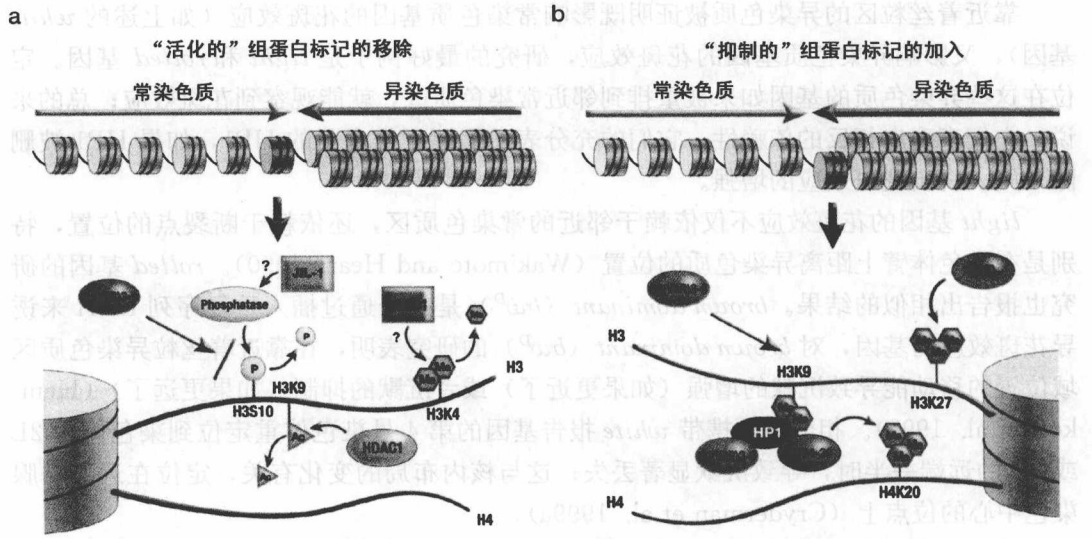



图 5-6　从常染色质状态向异染色质状态的转变需要一系列组蛋白修饰的改变

(a) 活化的基因被 H3K4me2 和 me3 标记；如果这个标记存在，那么在这个过程中一定会被 LSD1（在果蝇中尚未被鉴定）移除。H3K9 在常染色质中通常是乙酰化的；这个标记也一定会被 HDAC1（组蛋白去乙酰化酶）移除。H3S10 的磷酸化能够干扰 H3K9 的甲基化；去磷酸化似乎涉及一个 JIL-1 激酶羧基端相互作用的磷酸酶。(b) 这些转变设置了一个获得与基因沉默相关的组蛋白修饰的平台。SU(VAR)3-9甲基化 H3K9，HP1 结合到甲基化的 H3K9 上，随后募集 SUV4-20 甲基化 H4K20。E(Z)甲基化 H3K27 也可能会发生。

SU(VAR)3-9/HP1 形成复合体一样。

## 6. 果蝇异染色质是如何定位形成的

如前所述，尽管我们已经知道很多关于异染色质结构机制和生物化学方面的知识，但是有个问题始终没有解决，即在基因组正常构像中，异染色质形成是如何选择性地定位在某些区域的？所有的异染色质区域都有着一些共同的特点，其中有两个特点被认为是特定 DNA 序列异染色质组装的关键因素：核内空间上相对不同的异染色质区域的位置和 DNA 重复序列的存在。

总的来说，异染色质位于细胞核周边区域和围绕核仁的区域。在果蝇胚胎中，这种趋势尤为明显。异染色质在胚胎发生早期首次出现，这时细胞核移动到受精卵的周边区域。果蝇早期的发育是合胞体，一直到第 14 个核分裂周期，在核与核之间形成细胞壁，成为典型的囊胚。异染色质（着丝粒和第 4 号染色体）聚集在核内向卵外表面的一侧（Foe and Alberts 1985）。这种核内空间的划分在发育中一直存在，于是有了核内“异染色质空间”的概念。也许这些空间保持着异染色质形成所需的高浓度的因子（如 HP1 和 HKMT），而去除了常染色质组装和基因表达所需的因子，如组蛋白酰基转移酶（histone acetyl-transferase，HAT）和 RNA 聚合酶（RNA polymerase Ⅱ，RNA pol Ⅱ）。确实，靠近异染色质空间，不论是沿着染色体臂靠近还是在三维空间上靠近，都是 PEV 中的一个因素。

靠近着丝粒区的异染色质被证明既影响常染色质基因的花斑效应（如上述的 *white* 基因），又影响异染色质基因的花斑效应，研究的最好例子是 *light* 和 *rolled* 基因。定位在这些异染色质的基因如果被重排到邻近常染色质区，就能观察到花斑效应；总的来说，它们表现出相反的依赖性，它们的充分表达需要正常水平的 HP1，如果 HP1 被删除，则表现出花斑效应的增强。

*light* 基因的花斑效应不仅依赖于邻近的常染色质区，还依赖于断裂点的位置，特别是在染色体臂上距离异染色质的位置（Wakimoto and Hearn 1990）。*rolled* 基因的研究也报告出相似的结果。*brown dominant*（$bw^D$）是一个通过插入重复序列 DNA 来诱导花斑效应的基因，对 *brown dominant*（$bw^D$）的研究表明，在靠近着丝粒异染色质区域位置的移动能导致沉默的增强（如果更近了）或者沉默的抑制（如果更远了）（Henikoff et al. 1995）。相似地，携带 *white* 报告基因的第 4 号染色体重定位到染色体臂 2L 或 2R 的远端一半时，导致沉默显著丢失；这与核内布局的变化有关，定位在远离唾腺染色中心的位点上（Cryderman et al. 1999a）。

最近一项研究采用高分辨率显微镜检测了一个报告基因在同一个细胞中、在正常的表达时间框架内的基因活性（用针对基因产物特异的抗体）和在核内的定位（用荧光原位杂交）。一个 *white* 基因倒位的花斑效应 $bw^D$ 和 *lacZ* 转基因的花斑效应被用于研究分化的眼成虫盘或成虫的眼。这项研究发现基因的表达水平与报告基因在核内相对于臂间异染色质的位置密切相关，这支持了有“异染色质空间”存在的理论，并且基因在这个空间中的位置与基因沉默相关（Harmon and Sedat 2005）。然而，这种相关性并不是绝

对的。这并不奇怪，因为有一项关于 *white* 基因的研究表明，在小的第 4 号染色体（常靠近臂间异染色质块）上间隔分布着常染色质区和异染色质区。这些结果表明了其他的局部决定因子促进染色质包装成常染色质或异染色质。

在黑腹果蝇中，按照细胞生物学标准，有三分之一的基因组是异染色质。这包括围绕在着丝粒周围的大片染色质、端粒附近的一小片染色质、整个 Y 染色体和小的第 4 号染色体的大部分。着丝粒区 DNA 主要由这样两个部分组成：大段的（0.2～1Mb）卫星 DNA 和散布其中的复杂序列“岛”（主要是转座元件）（Le et al. 1995）。这个区域尽管基因贫乏，但并不是没有基因；现在估计有几百个基因位于臂间异染色质（Hoskins et al. 2002）。果蝇的端粒并不像其他地方一样有典型的富含 G 的重复序列，但却有多个拷贝的 HeT-A 和 TART 逆转座子。端粒相关序列（telomere-associated sequences，TAS）是一段段的 $10^2$～$10^3$ 核苷酸重复序列，只存在于近端。插入这个区域的 *white* 报告基因表现出花斑效应表型。尽管 Y 染色体携带有一些编码生育因子的基因，但是 Y 染色体的大部分是由卫星 DNA 组成的，并且在除雄性生殖系以外的细胞中始终保持浓缩状态。小的第 4 号染色体有 4.3Mb，其中有约 3Mb 由卫星 DNA 组成。远端的 1.2Mb 可以被认为是常染色质，它在唾腺中被多线化（图 5-4）；但由于其晚复制，完全缺乏有丝分裂交换以及它与 HP1、HP2 和 H3K9me2 相联系，又似乎是异染色质（图 5-4）。这个区域有 6～7 倍于常染色质臂的转座子片段的浓度，这和其他染色体上着丝粒异染色质与常染色质连接处的情况类似（Kaminker et al. 2002）。有趣的是，上面讨论的用 *white* 报告基因 *P* 元件对第 4 号染色体进行的一项研究中发现，常染色质（红眼表型）和异染色质（花斑眼睛表型）相间地分布在第 4 号染色体上（Sun et al. 2004）。

这个研究结果表明有能够驱动异染色质或常染色质形成的局部元件的存在。一个表型开关（从红到花斑或反之）的遗传学筛选证明，局部删除或加倍包围转座子报告基因的 5～80kb 的 DNA 能导致花斑效应的丢失或获得，这表明存在短程顺式作用的基因沉默决定子（图 5-7）。这个沉默依赖于 HP1 并且与染色质结构改变有关，因为核酸酶对染色质的可接触性改变了，说明染色质由常染色质向异染色质转变了。4 个区域中的一个区域的定位数据表明 *1360* 转座子是异染色质形成的一个靶点，而且在分散的重复元件中，一旦异染色质形成开始启动，就能沿着第 4 号染色体一直延伸 10kb，或者直到碰到来自常染色质决定子的竞争为止（Sun et al. 2004）。短程顺式作用的决定子与拷贝数有关，因为有研究发现报告基因 *P* 元件的正向或反向重复序列能够导致异染色质形成和基因沉默（Dorer and Henikoff 1994）。

DNA 中的这种顺式作用元件可能是与具有引发异染色质形成功能的蛋白质通过序列特异性结合而发挥作用的。与一些卫星 DNA 特异性结合的蛋白已经被鉴定出来（如 D1，Aulner et al. 2002）。这种相互作用的重要性从卫星 DNA 特异结合的药物对 PEV 的影响就可见一斑（Janssen et al. 2000）。然而，酵母以及植物中的一些研究结果（Elgin and Grewal 2003；Matzke and Birchler 2005；见第 8、9 章）提出了第二个模型，特别是一个以 RNAi 为基础的机制可以靶定重复元件导致异染色质形成。几个实验室的研究工作证明在果蝇中存在 RNAi 系统，且其通过转录后基因沉默（posttranscriptional

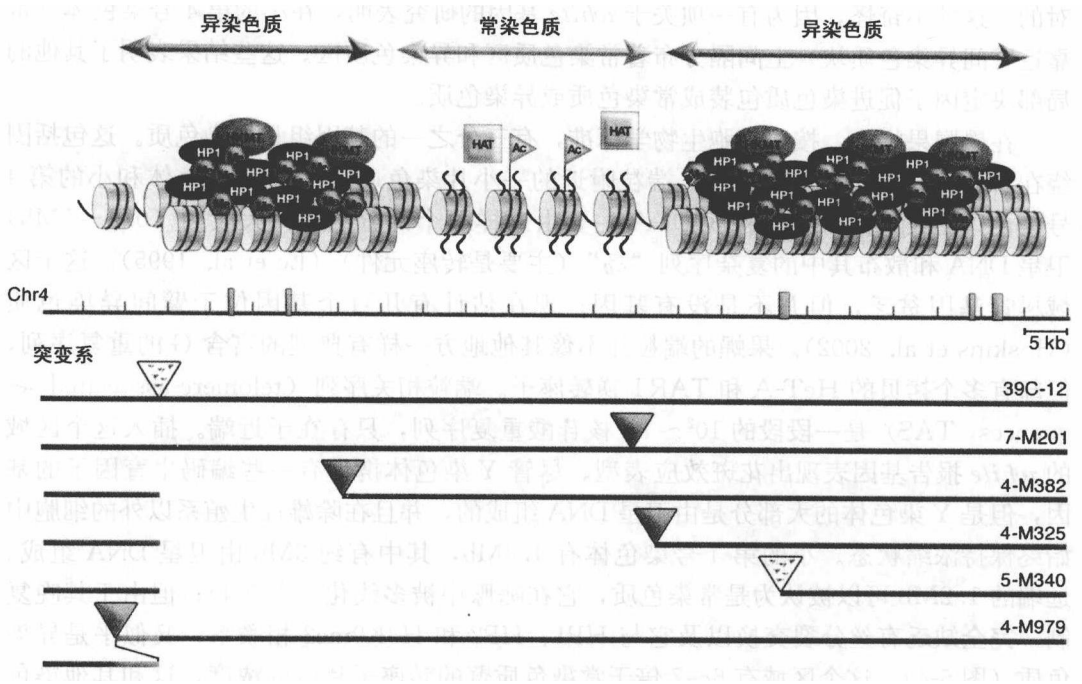



图 5-7 异染色质靶向的可能的模型

来自于重复序列的 dsRNA 被 RISC 加工成可能的“靶向复合体”，“靶向复合体”又在小 ssRNA 识别的位点指导组蛋白修饰或 HP1 的结合作为异染色质组装的起始步骤。第 4 号染色体的研究数据表明 DNA 转座子 *1360* 元件（橙色条）是异染色质形成的靶点之一；局部删除或加倍使 *P* 元件报告基因（三角）远离 *1360* 元件导致沉默的丢失（红三角表示红眼），而靠近 *1360* 元件则导致基因沉默（点三角表示花斑眼睛）（基于 Sun et al. 2004 的数据）。

gene silencing，PTGS）在发育调控中起到重要作用。黑腹果蝇中有两个基因编码 DICER 蛋白，有很多基因（*aubergine*，*agol*，*ago2*，*spindleE* [*aka homeless*]，*vasa intronic gene* [*VIG*]，*armitage*，*fmrl*）编码组装 RNA 诱导沉默复合体（RNAi-induced silencing complex，RISC）的组分（Sontheimer 2005）。这个系统被应用于重复序列的 PTGS，特别是串联重复的 *Stellate* 基因，几个逆转座子，乙醇脱氢酶（*Alcohol dehydrogenase*，*Adh*）转基因和在转录基因沉默（TGS）中的 *Adh* 转基因（Aravin et al. 2001；Pal-Bhadra et al. 2002）。在一个直接测试中，Pal-Bhadra 等（2004）发现 *piwi*（PAZ 结构域蛋白家族成员）和 *homeless*（一个 DEAD 盒螺旋酶）的突变抑制串联排列 *white* 基因相关的 PEV，并且发生在 *piwi*、*aubergine* 和 *homeless* 基因内的突变抑制位于臂间异染色质或第 4 号染色体的 *white* 转基因 P [hsp70-w] 的沉默。这个 PEV 的抑制与 H3K9 甲基化水平显著下降有关。在 40%已知的转座元件（包括 *1360*）和其他重复序列中，重复序列相关的小干扰 RNA（repeat-associated small interfering RNA，rasiRNA）已被鉴定出来了。

总结起来，上面讨论的结果表明异染色质形成也许既依赖核内的位置（也许提供充足的必需蛋白），也依赖于基于 RNAi 识别和来自重复序列特别是转座子的双链 RNA

处理的特异定位。这种通过 RISC 的定位能够或者带来 H3 甲基转移酶或者带来含有 HP1 的复合体（或者二者都有）并引发组装过程（见 5.4 节）。

## 7. 异染色质并不都是一样的

尽管上面描述异染色质都用的同样的术语，但是很显然，异染色质区在细节上是不同的。所有的异染色质区都特征性地富含重复序列 DNA，但重复序列 DNA 却并不相同，从着丝粒区串联排列地短重复序列（卫星 DNA）到高密度的分散重复序列（如见于第 4 号染色体的分散重复序列）。尽管所有的异染色质区似乎都与 HP1 和 H3K9me2 相关，但是显然涉及的蛋白复合体却在其他方面有着不同。检测 70 个不同修饰因子对不同的花斑基因（包括在臂间异染色质或 TAS 阵列中的 $w^{m4}$，$bw^{D}$，*P* 元件报告基因）的结果表明，尽管这些修饰因子的定位有着实实在在的重叠，但却同时有着惊人的复杂性。这些实验根据在既定空间里影响基因沉默的能力，将修饰子分为 7 个不同的组（Donaldson et al. 2002）。有趣的是，唯一影响在 TAS 阵列中沉默的修饰子是*Sun*(*var*) *3-9*的一个新的等位基因。

这些不同毫无疑问地反映了局部的生化变化或其中酶的变化。举个例子来说，细胞生物学结果表明尽管第 4 号染色体富集 H3K9me2，但是负责的酶却不是 SU(VAR) 3-9（Schotta et al. 2002；K. A. Haynes et al.，未发表）。甚至臂间异染色质，由于组成的 DNA 不同，从卫星 DNA 到成簇的分散重复序列也是镶嵌的（Le et al. 1995），可能利用不同的异染色质蛋白组合。这个结果可以在臂间异染色质不同部分对报告基因表达影响的研究中看到，在这个研究中，研究人员观察到花斑表型的严重程度不是简单地依赖顺式异染色质数量，还由局部异染色质环境的不同而不同（Howe et al. 1995）。异染色质相关的蛋白可能特异地在亚区域中起作用，其中包括 D1 蛋白，它倾向与 1.688g/cm$^3$ Ⅲ型卫星 DNA 结合（Aulner et al. 2002）；DDP1，一个与 vigilin 同源的多 KH 结构域蛋白，倾向与 dodeca 卫星 DNA 富含嘧啶 C 链结合（Cortes and Azorin 2000）。

## 8. 不同生物中的 PEV、异染色质形成和基因沉默

位置效应花斑现象最早在果蝇中被发现，主要是因为果蝇是第一个被用于 X 射线诱变的生物。X 射线比其他常用的诱变剂更能诱导染色体重排，进而导致 PEV。在小鼠中分离到类似的突变，如花斑皮色显示出 PEV。遗传分析表明一个携带皮色野生型基因的常染色体区域被插入到 X 染色体中（Cattanach 1961；Russel and Bangham 1961）。花斑效应只能在携带这个插入片段且原皮色基因突变纯合子雌性鼠中观察到。在这些雌性鼠中，由于 X 染色体异染色质化失活，野生型等位基因也被沉默了（见第 17 章）。在植物中，唯一被描述过的明确的 PEV 的例子是 *Oenothera blandina*（Catcheside 1939）。和果蝇一样，这些例子中的 PEV 沉默常染色质基因都与基因被置于新的异染色质附近有关。

转录水平基因沉默（repeat-induced gene silencing，RIGS，重复序列诱导的基因沉默）也在重复序列中发现过，尤其是在植物中。对受影响序列的分析表明它们与异染色质和 PEV 沉默的区域有着相似的表观遗传标记（组蛋白和 DNA 甲基化）。如果将含有荧光酶基因串联排列的 DNA 片段导入到拟南芥中，可以看到荧光酶基因表达的花斑效应。在这里，观察到的基因沉默也是源于异染色质形成。它们的分子机制在高等真核生物中是保守的。

果蝇中异染色质基因沉默的一个中心特点就是 HP1 与 H3K9me2 和 SU(VAR)3-9 HKMT 的相互作用。HP1 从 *S. pombe* 到人都是保守的，并且结合在臂间异染色质上。人 HP1 基因可以弥补果蝇 HP1 缺陷（Ma et al. 2001）。然而，HP1 还没有在植物中被鉴定过。SU(VAR)3-9存在更为广泛，在酿酒酵母（Clr4）、红色链孢霉菌（DIM5）、拟南芥和哺乳动物（SUV39H）中都已鉴定。所有的 SU(VAR)3-9同源物都催化 H3K9 的甲基化并在异染色质形成中起作用。人 SUV39H1 转基因也能弥补果蝇中内源同源蛋白的缺失（Scotta et al. 2002）。在高等植物（水稻、拟南芥、玉米）中，几个 SU(VAR)3-9同源蛋白（SU(VAR)3-9 homologous protein，SUVH）也被发现了（Baumbusch et al. 2001）。大量的 HKMT 也许反映了植物发育中的可塑性或者对环境因子响应的需要（进一步讨论见第 9 章）。4 个 SUVH 蛋白——SUVH1、SUVH2、SUVH4（KYP）和 SUVH6 都被详细地研究过（Jackson et al. 2002；Naumann et al. 2005）。所有这些都是 H3K9 甲基转移酶。SUVH2 在控制异染色质状态中起到关键作用，对异染色质形成表现出的剂量依赖效应与果蝇中的 SU(VAR)3-9类似（Naumann et al. 2005）。SUVH2 的功能丧失显著抑制重复序列依赖的基因沉默，而过表达则增强了荧光酶基因在植物中的沉默。

在果蝇 *Su(var)*突变中鉴定出来的其他基因编码的蛋白都有着保守的功能。SUV4-20 HKMT 在哺乳动物和在果蝇中（Schotta et al. 2004）控制着 H4K20 的三甲基化。组蛋白去甲基化酶、乙酰化酶和去乙酰化酶都是保守的（T. Rudolph et al.，未发表）许多控制组蛋白修饰的关键酶的进化保守性正支持了组蛋白密码学说（Jenuwein and Allis 2001）。然而，在果蝇、哺乳动物和植物（拟南芥）中研究异染色质特异的组蛋白修饰标记时，也鉴定出一些种属特异的元件。

哺乳动物中组成型异染色质的显著标记包括 H3K9me3、H3K27me1 和 H4K20me3（Peters et al. 2003；Schcotta et al. 2004）。果蝇异染色质的特征包括 H3K9me1/me2、H3K27me1/me2/me3 和 H4K20me3（Schotta et al. 2002 2004；Ebert et al. 2004）。与哺乳动物相反，果蝇中 H3K9me3 不足。在哺乳动物中，H3K9me1 不是异染色质标记。在拟南芥中，与果蝇一样，H3K9me1/me2 是异染色质标记，而 H3K9me3 是常染色质标记（Naumann et al. 2005）。H3K27me1 和 H3K27me2 在拟南芥中是异染色质标记，而在果蝇常染色质和异染色质中都有这两个标记。H3K27me3 在拟南芥中是常染色质特异的标记。在拟南芥中 H4K20me1 是异染色质标记，而 H4K20me2 和 H4K20me3 是常染色质标记。拟南芥和动物另一个显著的区别是 H3S10 在染色体上的分布。这个标记在拟南芥中是异染色质标记（A. Fischer and G. Reuter，未发表）而在果蝇中是常染色质标记（Wang et al. 2001；Ebert et al. 2004）。哺乳动物、果蝇和拟南芥在异染色

质特异的组蛋白修饰标记上的相似性和差异性清楚地表明，组蛋白密码不是通用的，而是以不同的“方言”形式存在。

## 9. 总结：关于异染色质我们知之甚少

尽管 PEV 为我们研究异染色质形成和基因沉默提供了一个极佳的机会，但是其表型本身依然令人感到迷惑。为什么我们观察到基因沉默花斑模式？是什么破坏了平衡，导致染色质由活跃状态向沉默状态转换，或者相反的转换？为什么它似乎在同源细胞中遗传？PEV 是作为维持报告基因“开”或“关”的问题来研究的，但是在许多例子中（特别是当我们使用了基于 *P* 元件的报告基因），研究人员能够在黄色或浅黄的背景下看到小红点，这表明基因表达被一致地下调了，但是在有些细胞中这种下调丢失了。对此果蝇系进行的分析也许能够发现对基因表达影响呈中间状态的染色质状态。支持基因沉默丢失和维持的初步模型研究是基于这种总体行为的，但是最终的模型必定是复杂的，涉及众多相互作用的蛋白（Henikoff 1996）。也有人认为核小体是个总体装置，综合所有的组蛋白修饰，呈现出特殊的蛋白结合模式并促进这个区域重塑。染色质状态也许反映了不同组蛋白修饰状态相互竞争的结果。这样的模型也许在上述的作用的分类上非常有用。有结果证明 GAL4 依赖的报告基因沉默的频率对 GAL4 敏感（Ahmad and Henikoff 2001），上述模型与这个结果也是一致的。

RNAi 系统为异染色质形成的定位提供了一个合理的机制，这可能是通过定位一个蛋白复合体来实现的，这个复合体包含 HP1、H3K9 HKMT 或者两个都有。然而，很多问题依然存在。dsRNA 从哪儿来？它一定是顺式产生的吗（如 *S. pombe* 中的研究结果），或它起反式作用吗（如植物中的研究结果）；即一个 *1360* 元件的产物能够使所有的 *1360* 位点定位形成异染色质吗？是不是所有的重复序列元件都有这种潜在的定位功能？从第 4 号染色体的分析来看，这似乎是不可能的。如果只是一小部分的重复序列元件起到关键作用，那么又是什么决定了这个选择呢？第 4 号染色体的研究结果表明关键在于重复序列元件的密度和分布以及附近基因的表达。这需要对基因组进行测序以便确认这个特点。

异染色质蔓延是如何完成的，这种蔓延的壁垒是什么？在果蝇中并没有关于蔓延 RNAi 的证据，即沉默没有蔓延到 dsRNA 序列上游转录物（Celotto and Gravely 2002）。这与在这个系统中没有发现 RNA 依赖的 RNA 聚合酶的结果相一致。基于 HP1、H3K9me2 和一个 HKMT 的组装系统也许很好地解释了异染色质能蔓延近 10kb 的实验结果，如在第 4 号染色体上观察到的结果；这个类型的蔓延可能会被某一个位点组蛋白的乙酰化所限制。但在重排中发生的异染色质蔓延又是怎么回事呢？这种蔓延能够持续几十万个碱基。这种蔓延的不连续性，似乎关键依赖于染色质蛋白，特别是 JIL-1，它以不依赖其激酶活性的方式起作用。除了这些，还有其他很多问题有待解决。

## 致 谢

我们感谢 Gabriella Farkas 为本章所作的插图，感谢 Anja Ebert 提供免疫组化照片，感谢我们研究组成员对本章节的严格审阅，以及德意志联合研究会和欧盟表观基因组学先进网络的支持及美国国立卫生研究院的基金资助。

张道永 译
王大亮 校
孙方霖 审校

## 参考文献


Ahmad K. and Henikoff S. 2001. Modulation of a transcription factor counteracts heterochromatin gene silencing in *Drosophila*. *Cell* **104:** 839–847.

Aravin A.A., Numova H. M., Tulin A. V., Vagin, V.V., Rozovsky Y.M., and Gvozdev V.A. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11:** 1017–1027.

Aravin A. A., Lagos-Quintana M., Yalcin A., Zavolan M., Marks D., Snyder B., Gaasterland T., Meyer J., and Tuschl T. 2003. The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell* **5:** 337–350.

Ashburner M., Golic K.G., and Hawley R.S. 2005a. Chromosomes. In *Drosophila: A laboratory handbook*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp. 24–57.

———. 2005b. Position effect variegation. In *Drosophila: A laboratory handbook*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, pp.1007–1049.

Aulner N., Monod C., Mandicourt G., Jullien D., Cuvier O., Sall A., Janssen S., Laemmli U.K, and Kas E. 2002. The AT-hook protein D1 is essential for *Drosophila melanogaster* development and is implicated in position-effect variegation. *Mol. Cell. Biol.* **22:** 1218–1232.

Baksa K., Morawietz H., Dombradi V., Axton M., Taubert H., Szabo G., Török I., Gyurkovics H., Szöör B., Gloover D., et al. 1993. Mutations in the phosphatase 1 gene at 87B can differentially affect suppression of position-effect variegation and mitosis in *Drosophila melanogaster*. *Genetics* **135:** 117–125.

Bannister A.J., Zegermann P., Patridge J.F., Miska E.A., Thomas J.O., Allshire T.C., and Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120–124.

Baumbusch L.O., Thorstensen T., Krauss V., Fischer A., Naumann K., Assalkhou R., Schulz I., Reuter G., and Aalen R. 2001. The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionary conserved classes. *Nucleic Acids Res.* **29:** 4319–4333.

Belyaeva E.S., Demakova O.V., Umbetova G.H., and Zhimulev I.F. 1993. Cytogenetic and molecular aspects of position-effect variegation in *Drosophila melanogaster*. V. Heterochromatin-associated protein HP1 appears in euchromatic chromosomal regions that are inactivated as a result of position-effect variegation. *Chromosoma* **102:** 583–590.

Bender W., Spierer P., and Hogness D.S. 1983. Chromosome walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the bithorax complex of *Drosophila melanogaster*. *J. Mol. Biol.* **168:** 17–33.

Birchler J.A., Bhadra U., Rabinow L., Linsk R., and Nguyen-Huyuh A.T. 1994. *Weakener of white* (*Wow*), a gene that modifies the expression of white eye color locus and that suppresses position effect variegation in *Drosophila melanogaster*. *Genetics* **137:** 1057–1070.

Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16:** 6–21.

Bushey D. and Locke J. 2004. Mutations in *Su(var)205* and *Su(var)3-7* suppress P-element-dependent silencing in *Drosophila melanogaster*. *Genetics* **168:** 1395–1411.

Catcheside D.G. 1939. A position effect in *Oenothera*. *J. Genet.* **38:** 345–352.

Cattanach B.M. 1961. A chemically-induced variegated-type position effect in the mouse. *Z. Vererbungsl.* **92:** 165–182.

Celotto A.M. and Graveley B.R. 2002 Exon-specific RNAi: A tool for dissecting the functional relevance of alternative splicing. *RNA* **8:** 718–724.

Cleard F. and Spierer P. 2001. Position-effect variegation in *Drosophila*: The modifier *Su(var)3-7* is a modular DNA-binding protein. *EMBO Rep.* **21:** 1095–1100.

Cleard F., Delattre M., and Spierer P. 1997. SU(VAR)3-7 a *Drosophila* heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. *EMBO J.* **16:** 5280–5288.

Cooper K.W. 1959. Cytogenetic analysis of major heterochromatic elements (especially Xh and Y) in *Drosophila melanogaster* and the theory of "heterochromatin". *Chromosoma* **10:** 535–588.

Cortes A. and Azorin F. 2000. DDP1, a heterochromatin-asociated multi-KH-domain protein of *Drosophila melanogaster*, interacts specifically with centromeric satellite DNA sequences. *Mol. Cell. Biol.* **20:** 3860–3869.

Cryderman D.E., Morris E.J., Biessmann H., Elgin S.C.R., and Wallrath L.L. 1999a. Silencing at *Drosophila* telomeres: Nuclear organization and chromatin structure play critical roles. *EMBO J.* **18:** 3724–3735.

Cryderman D. E., Tang H., Bell C., Gilmour D.S., and Wallrath L.L. 1999b. Heterochromatic silencing of *Drosophila* heat shock genes acts at the level of promoter potentiation. *Nucleic Acids Res.* **27:** 3364–3370.

Czermin B., Schotta G., Hülsmann B.B., Brehm A., Becker P.B., Reuter G., and Imhof A. 2001. Physical and functional interaction of SU(VAR)3-9 and HDAC1 in *Drosophila*. *EMBO Rep.* **2:** 915–919.

Delattre M., Spierer A., Tonka C.H., and Spierer P. 2000. The genomic silencing of position-effect variegation in *Drosophila melanogaster*: Interaction between the heterochromatin-associated proteins Su(var)3-7 and HP1. *J. Cell Sci.* **113:** 4253–4261.

Demerec M. and Slizynska H. 1937. Mottled white 258-18 of *Drosophila melanogaster*. *Genetics* **22:** 641–649.

Donaldson K.M., Lui A., and Karpen G.H. 2002. Modifiers of terminal deficiency-associated position effect variegation in *Drosophila*. *Genetics* **160:** 995–1009.

Dorer D.R. and Henikoff S. 1994. Expansion of transgene repeats causes heterochromatin formation and gene silencing in *Drosophila*. *Cell* **77:** 993–1002.

Dorn R., Krauss V., Reuter G., and Saumweber H. 1993a. The enhancer

of position-effect variegation *E(var)93D*, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. *Proc. Natl. Acad. Sci.* **90:** 11376–11380.

Dorn R., Szidonya J., Korge G., Sehnert M., Taubert H., Archoukieh I., Tschiersch B., Morawietz H., Wustmann G., Hoffmann G., and Reuter G. 1993b. P Transposon-induced dominant enhancer mutations of position-effect variegation in *Drosophila melanogaster*. *Genetics* **133:** 279–290.

Ebert A., Schotta G., Lein S., Kubicek S., Krauss V., Jenuwein T., and Reuter G. 2004. Su(var) genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. *Genes Dev.* **18:** 2973–2983.

Eissenberg J.C., Morris G.D., Reuter G., and Hartnett T. 1992. The hetero-chromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics* **131:** 345–352.

Eissenberg J.C., James T.C., Foster-Hartnett D.M., Hartnett T., Ngan V., and Elgin S.C.R. 1990. A mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position effect variegation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* **87:** 9923–9927.

Elgin S.C.R. and Grewal S.I.S. 2003. Heterochromatin: Silence is Golden. *Curr. Biol.* **13:** R895–R898.

Fanti L. and Pimpinelli S. 2004. Immunostaining of squash preparations of chromosomes of larval brains. *Methods Mol. Biol.* **247:** 353–361.

Fanti L., Giovinazzo G., Berloco M., and Pimpinelli S. 1998. The heterochromatin protein 1 prevents telomere fusions in *Drosophila*. *Mol. Cell* **2.** 527–538.

Farkas G., Gausz J., Galloni M., Reuter G., Gyurkovics H., and Krach F. 1994. The *trithorax-like* gene encodes the *Drosophila* GAGA factor. *Nature* **371:** 806–808.

Fischle W., Wang Y., Jacobs S.A., Kim Y., Allis C.D., and Khorasanizadeh S. 2003. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* **17:** 1870–1881.

Flybase 2005. The *Drosophila* database. Available from World Wide Web at the URLs http://morgan/harvard.edu and http://www.ebi.ac.uk/flybase/

Foe V.E. and Alberts B.M. 1985. Reversible chromosome condensation induced in *Drosophila* embryos by anoxia: Visualization of interphase nuclear organization. *J. Cell Biol.* **100:** 1623–1636.

Harmon B. and Sedat J. 2005. Cell-by-cell dissection of gene expression and chromosomal interactions reveals consequences of nuclear reorganization. *PLoS Biol.* **3:** e67.

Hartmann-Goldstein I.J. 1967. On the relationship between heterochromatization and variegation in *Drosophila*, with special reference to temperature sensitive periods. *Genet. Res.* **10:** 143–159.

Henikoff S. 1996. Dosage-dependent modification of position-effect variegation in *Drosophila*. *BioEssays* **18:** 401–409.

Henikoff S., Jackson J.M., and Talbert P.B. 1995. Distance and pairing effects on the *brown*$^{Dominant}$ heterochromatic element in *Drosophila*. *Genetics* **140:** 1007–1017.

Hoskins R.A., Smith C.D., Carlson J.W., Carvalho A.B., Halpern A., Kaminker J.S., Kennedey C., Mungall C.J., Sullivan B.A., Sutton G.G., et al. 2002. Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. *Genome Biol.* **3:** RESEARCH0085.

Howe M., Dimitri P., Berloco M., and Wakimoto B.T. 1995. *cis*-effects of heterochromatin on heterochromatic and euchromatic gene activity in *Drosophila melanogaster*. *Genetics.* **140:** 1033–1045.

Jackson J.P., Lindroth A.M., Cao X., and Jacobsen S.E. 2002. Control of CpNpG DNA methylation by the *KRYPONITE* histone H3 methyltransferase. *Nature* **416:** 556–560.

James T.C. and Elgin S.C.R. 1986. Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. *Mol. Cell Biol.* **6:** 3862–3872.

James T.C., Eissenberg J.C., Craig C., Dietrich V., Hobson A., and Elgin S.C.R. 1989. Distribution patterns of HP1, a heterochromatin-associated nonhistone chromosomal protein of *Drosophila*. *Eur. J. Cell Biol.* **50:** 170–180.

Janssen S., Cuvier O., Muller M., and Laemmli U.K. 2000. Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to *Drosophila melanogaster*. *Mol. Cell* **6:** 1013–1024.

Jenuwein T. and Allis C.D. 2001. Translating the histone code. *Science* **293:** 1074–1080.

Jones R.S. and Gelbart W.M. 1993. The *Drosophila* Polycomb-group gene *Enhancer of zeste* contains a region with sequence similarity to trithorax. *Mol. Cell. Biol.* **13:** 6357–6366.

Kaminker J.S. Bergman C.M., Kronmiller B., Carlson J., Svirskas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., et al. 2002. The transposable elements of the *Drosophila melanogaster* genome: A genomics perspective. *Genome Biol.* **3:** RESEARCH0084.

Karpen G.H. and Spradling A.C. 1990. Reduced DNA polytenization of a minichromosome region undergoing position-effect variegation in *Drosophila*. *Cell* **63:** 97–107.

Krauss V. and Reuter G. 2000. Two genes become one: The genes encoding heterochromatin protein SU(VAR)3-9 and translation initiation factor subunit eIF-2γ are joined to a dicistronic unit in holometabolic insects. *Genetics* **156:** 1157–1167.

Kunert N., Marhold J., Stanke J., Stach D., and Lyko F. 2003. A Dnmt2-like protein mediates DNA methylation in *Drosophila*. *Development* **130:** 5083–5090.

Lachner M., O'Carroll D., Rea S., Mechtler K., and Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410:** 116–120.

Laible G., Wolf A., Dorn R., Reuter G., Nislow C., Lebesorger A., Popkin D., Pillus L., and Jenuwein T. 1997. Mammalian homologues of the Polycomb-group gene *Enhancer of zeste* mediate gene silencing in *Drosophila* heterochromatin and at *S. cerevisiae*. *EMBO J.* **16:** 3219–3232.

Larsson J., Zhang J., and Rasmuson-Lestander A. 1996. Mutations in the *Drosophila melanogaster* S-adenosylmethionine synthase suppress position-effect variegation. *Genetics* **143:** 887–896.

Le M.H., Duricka D., and Karpen G.H. 1995. Islands of complex DNA are widespread in *Drosophila* centric heterochromatin. *Genetics* **141:** 283–303.

Locke J., Kotarski M. A., and Tartof K.D. 1988. Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* **120:** 181–198.

Lu B.Y., Bishop C.P., and Eissenberg J.C. 1996. Developmetal timing and tissue specificity of heterochromatin-mediated silencing. *EMBO J.* **15:** 1323–1332.

Ma J., Hwang K.K., Worman H.J., Courvalin J.C., and Eissenberg J.C. 2001. Expression and functional analysis of three isoforms of human heterochromatin-associated protein HP1 in *Drosophila*. *Chromosoma* **109:** 536–544.

Martienssen R.A. and Colot V. 2001. DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science* **293:** 1070–1074.

Mason J.M., Ransom J., and Konev A.Y. 2004. A deficiency screen for dominant suppressors of telomeric silencing in *Drosophila*. *Genetics* **168:** 1353–1370.

Matzke M.A. and Birchler J.A. 2005. RNAi-mediated pathways in the nucleus. *Nat. Rev. Genet.* **6:** 24–35.

Mottus R., Sobels R.E., and Grigliatti T.A. 2000. Mutational analysis of a histone deacetylase in *Drosophila melanogaster*: Missense mutations suppress gene silencing associated with position effect variegation. *Genetics* **154:** 657–668.

Muller H.J. 1930. Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* **22:** 299–334.

Naumann K., Fischer A., Hofmann I., Krauss V., Phalke S., Irmler K., Hause G., Aurich A.C., Dorn R., Jenuwein T., and Reuter G. 2005. Pivotal role of *At*SUVH2 in control of heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J.* **24:** 1418–1429.

Pal-Bhadra M., Bhadra U., and Birchler J.A. 2002. RNAi related mechanisms affect both transcriptional and post-transcriptional transgene silencing in *Drosophila*. *Mol. Cell* **9:** 315–327.

Pal-Bhadra M., Leibovitch B.A., Gandhi S.G., Rao M., Bhadra U., Birchler J.A., and Elgin S.C.R. 2004. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303:** 669–672.

Paro R. and Hogness D.S. 1991. The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl. Acad. Sci.* **88:** 263–267.

Peters A.H.F.M., Kubicek S., Mechtler K., O'Sullivan J., Derijck A.A.H.A., Perez-Burgos L., Kohlmaier A., Opravil S., Tachibana M., Shinkai Y., et al. 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* **12:** 1577–1589.

Platero J.S., Sharp E.J., Adler P.N., and Eissenberg J.C. 1996. In vivo assay for protein-protein interaction using *Drosophila* chromosomes. *Chromosoma* **104:** 393–404.

Rea S., Eisenhaber F., O'Carroll D., Strahl B.D., Sun Z-W., Schmid M., Opravil S., Mechtler K., Ponting C.P., Allis C.D., and Jenuwein T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406:** 593–599.

Reuter G., Werner W., and Hofmann H.J. 1982. Mutants affecting position-effect heterochromatinization in *Drosophila melanogaster*. *Chromosoma* **85:** 539–551.

Reuter G., Wolff I., and Friede B. 1985. Functional properties of the heterochromatic sequences inducing $w^{m4}$ position-effect variegation in *Drosophila melanogaster*. *Chromosoma* **93:** 132–139.

Reuter G., Giarre N., Farah J., Gausz J., Spierer A., and Spierer P. 1990. Dependence of position-effect variegation in *Drosophila* on dose of a gene encoding an unusual zinc-finger protein. *Nature* **344:** 219–223.

Rice J.C., Briggs S.D., Ueberheide B., Barber C.M., Shabanowitz J., Hunt D.F., Shinkai Y., and Allis C.D. 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell* **12:** 1591–1598.

Richards E.J. and Elgin S.C.R. 2002. Epigenetic codes for heterochromatin formation and silencing: Rounding up the usual suspects. *Cell* **108:** 489–500.

Russel L.B. and Bangham J.W. 1961. Variegated type position effects in the mouse. *Genetics* **46:** 509–525.

Schotta G., Ebert A., Krauss V., Fischer A., Hoffmann J., Rea S., Jenuwein T., and Reuter G. 2002. Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J.* **21:** 1121–1131.

Schotta G., Ebert A., Dorn R., and Reuter G. 2003. Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Semin. Cell Dev. Biol.* **14:** 67–75.

Schotta G., Lachner M., Sarma K., Ebert A., Sengupta R., Reuter G., Reinberg D., and Jenuwein T. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **18:** 1251–1262.

Schultz J. 1950. Interrelations of factors affecting heterochromatin-induced variegation in *Drosophila*. *Genetics* **35:** 134.

Seum C., Pauli D., Delattre M., Jaquet Y., Spierer A., and Spierer P. 2002. Isolation of *Su(var)3-7* mutations by homologous recombination in *Drosophila melanogaster*. *Genetics* **161:** 1125–1136.

Shaffer C.D., Stephens G.E., Thompson B.A., Funches L., Bernat J.A., Craig C.A., and Elgin S.C.R. 2002. Heterochromatin protein 2 (HP2), a partner of HP1 in *Drosophila* heterochromatin. *Proc. Natl. Acad. Sci.* **99:** 14332–14337.

Shi Y., Lan F., Matson C., Mulligan P., Whetstine J.R., Cole P.A., Casero R.A., and Shi Y. 2005. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119:** 941–953.

Silver L.M., and Elgin S.C.R. 1976. A method for determination of the *in situ* distribution of chromosomal proteins. *Proc. Natl. Acad. Sci.* **73:** 423–427.

Sinclair D.A.R., Mottus R.C., and Grigliatti T.A. 1983. Genes which suppress position effect variegation in *Drosophila melanogaster* are clustered. *Mol. Gen. Genet.* **191:** 326–333.

Sinclair D.A.R., Clegg N.J, Antonchuk J, Milner T.A., Stankunas K., Ruse C., Grigliatti T.A., Kassis J., and Brock H.W. 1998. *Enhancer of Polycomb* is a suppressor of position-effect variegation in *Drosophila melanogaster*. *Genetics* **148:** 211–220.

Sontheimer E.J. 2005. Assembly and function of RNA silencing complexes. *Nat. Rev. Mol. Cell Biol.* **6:** 127–138.

Spofford J.B. 1967. Single-locus modification of position-effect variegation in *Drosophila melanogaster*. I. white variegation. *Genetics* **57:** 751–766.

Stephens G.E., Craig C.A., Li Y., Wallrath L.L., and Elgin S.C.R. 2003. Immunofluorescent staining of polytene chromosomes: Exploiting genetic tools. *Methods Enzymol.* **376:** 372–393.

Sun F.-L., Cuaycong M.H., and Elgin S.C.R. 2001. Long-range nucleosome ordering is associated with gene silencing in *Drosophila melanogaster* pericentromeric heterochromatin. *Mol. Cell. Biol.* **21:** 2867–2879.

Sun F.-L., Haynes K., Simpson C.L., Lee S.D., Collins L., Wuller J., Eissenberg J.C., and Elgin S.C.R. 2004. *cis*-acting determinants of heterochromatic formation on *Drosophila melanogaster* chromosome four. *Mol. Cell. Biol.* **24:** 8210–8220.

Tartof K.D., Hobbs C., and Johnes M. 1984. A structural basis of variegating position effects. *Cell* **37:** 869–878.

Tschiersch B., Hofmann A., Krauss V., Dorn R., Korge G., and Reuter G. 1994. The protein encoded by the *Drosophila* position effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* **13:** 3822–3831.

Wakimoto B.T. and Hearn M.G. 1990. The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2L of *Drosophila melanogaster*. *Genetics* **125:** 141–154.

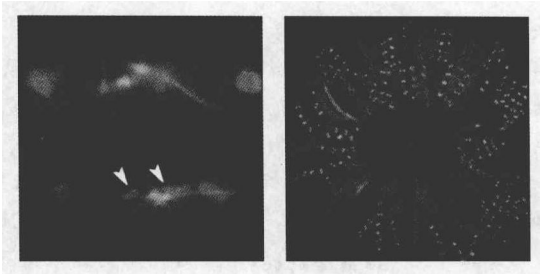
Wallrath L.L. and Elgin S.C.R. 1995. Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev.* **9:** 1263–1277.

Wallrath L.L., Gunter V.P., Rosman L.E., and Elgin S.C.R. 1996. DNA representation of variegating heterochromatic P element inserts in diploid and polytene tissue of *Drosophila melanogaster*. *Chromosoma* **104:** 519–527.

Wang Y., Zhang W., Jin Y., Johansen J., and Johansen K.M. 2001. The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. *Cell* **105:** 433–443.

Zhimulev I.F., Belyaeva E.S., Fomina O.V., Protopopov M.O., and Bolshkov V.N. 1986. Cytogenetic and molecular aspects of position effect variegation in *Drosophila melanogaster*. *Chromosoma* **94:** 492–504.

# 第 6 章



# 表观遗传学研究的真菌模型：裂殖酵母和链孢霉

Robin C. Allshire[1] and Eric U. Selker[2]

[1] *Wellcome Trust Centre for Cell Biology, The University of Edinburgh, Edinburgh EH9 3JR, Scotland, United Kingdom*

[2] *Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403-1229*

## 本章目录

1. 裂殖酵母
   - 1.1 裂殖酵母中的染色质沉默与酿酒酵母有所不同
   - 1.2 裂殖酵母着丝粒处的基因是沉默的
   - 1.3 裂殖酵母着丝粒由异染色质和中央动粒结构组成
   - 1.4 着丝粒外侧重复序列本身导致了沉默染色质的装配
   - 1.5 RNA 干扰指导沉默染色质装配
   - 1.6 RNA 聚合酶Ⅱ合成的着丝粒重复转录产物使 RNAi 和染色质修饰相联系
   - 1.7 着丝粒沉默染色质参与姐妹染色体着丝粒黏联和染色体正常分离
   - 1.8 功能性着丝粒表观遗传学状态的继承
   - 1.9 真菌中沉默机制的多样性
2. 链孢霉：历史和特性
   - 2.1 链孢霉中的 DNA 甲基化
   - 2.2 RIP——既涉及遗传学机制也涉及表观遗传学机制的基因组防卫系统
   - 2.3 研究 RIP 的遗迹有助于了解 DNA 甲基化的调控
   - 2.4 平抑
   - 2.5 非配对 DNA 导致的减数分裂沉默
   - 2.6 RIP、平抑和 MSUD 可能的功能和实际应用
3. 总结

致谢

参考文献

## 概要

真菌为研究转录活跃区域（常染色质）和转录沉默区域（异染色质）中染色质的结构和功能提供了一个良好模型。酿酒酵母（酿酒酵母）已经成为研究常染色质区与转录相关的染色质结构的珍贵模型，并为研究沉默染色质提供了一个极好的范例。另一方面，作为研究沉默染色质的工具，裂殖酵母和链孢霉（一种丝状真菌）与高等真核生物更为接近。在这些真菌中，异染色质区域相对较小并且不为生存所必需，这使得可以对其进行分析和操作。我们对着丝粒周围和端粒区域的异染色质的了解，在酿酒酵母和裂殖酵母中最为详尽，然而裂殖酵母染色质沉默的机制表现出与高等真核生物中保守的特征。与酿酒酵母相比，我们这章要讨论的真菌（链孢霉和裂殖酵母）利用RNA干扰（RNAi）、组蛋白H3K9甲基化和类似异染色质蛋白1（heterochromatin protein 1，HP1）的蛋白来产生沉默染色质，这些方式在植物和多细胞动物中保守或相似。另外，链孢霉中存在DNA甲基化，这在许多高等真核生物中是异染色质的典型特征，也是一种经典的表观遗传现象。首先，在简要介绍裂殖酵母生活习性之后，我们将讨论其异染色质的性质和功能。然后，我们转入讨论链孢霉，描述这种丝状真菌是如何对表观遗传学研究作出贡献的。

## 1. 裂殖酵母

裂殖酵母（*S. pombe*）是在亚热带地区啤酒生产的发酵过程中被发现的；“pombe”在斯瓦希里语中是啤酒的意思。裂殖酵母主要以单倍体单细胞形式存在。在营养丰富的培养基中，野生型细胞大约每两小时经历一次有丝分裂。但是某些条件或者条件突变体，可以被用来使细胞周期停止于某个阶段或使细胞同步化于 $G_1$/S 期、$G_2$ 期或者 M 期。这很有用，因为在快速生长的培养物中 $G_1$ 期很短，细胞经过分裂后几乎立即进入 S 期；细胞周期的大部分时间处于 $G_2$ 期（图 6-1）（Egel 2004）。

和酿酒酵母一样，裂殖酵母可以在相反的接合型之间转换，被命名为正（＋）或负（－）。接合型等同于高等真核生物中的雌雄两性，尽管他们是单倍体。两种接合型信息分别存在于基因组中被表观遗传机制调控的沉默盒子中 mat2-P（＋）或 mat3-M（－）位点。两个沉默区域只为接合型的确定提供遗传模板，接合型本身由有活性的 *mat1* 位点中的＋或－信息所决定。活性 *mat1* 位点信息的转换，是通过沉默位点和 mat1 位点精确的重组来实现的（Egel 2004）。当氮源不足时，细胞停止分裂停滞于 $G_1$ 期，这会促使细胞进入生活周期中的有性生殖阶段，＋和－的细胞接合形成二倍体合子（图 6-1）。在接合与核融合之后，发生减数分裂前 DNA 复制（DNA 量从 2N 变为 4N），然后发生同源染色体的配对和重组，接着是减数分裂的Ⅰ期和Ⅱ期。于是产生包在子囊中的四个单倍体孢子。之后当氮源丰富的时候孢子会萌发，恢复无性生殖与有丝分裂。

不能发生接合型转换的菌株已经被分离或者构建，它们中所有的细胞都属于同一种接合型。这有利于控制两种不同基因型的菌株间的交配。尽管裂殖酵母一般是单倍体，但也有可能筛选到二倍体的菌株。这种二倍体菌株以无性有丝分裂的方式生长，当氮缺乏时，它们也经历减数分裂然后形成“非合子子囊（azygotic asci）”。

### 1.1　裂殖酵母中的染色质沉默与酿酒酵母有所不同

在研究沉默染色质和相关的表观遗传效应时，裂殖酵母是非常有用的模型。在通过形成异染色质来导致沉默的机制上，它不同于酿酒酵母，而是更加类似于链孢霉、植物和多细胞动物。端粒和接合型位点在酿酒酵母和裂殖酵母中都是沉默的，但在裂殖酵母中，着丝粒也表现出沉默的特征。就如在第四章中所描述的，酿酒酵母使用一类独特的蛋白——Sir 蛋白来沉默染色质。与此不同，裂殖酵母和其他的真核生物利用不同的染色质修饰（特别是组蛋白 H3K9 的甲基化）与 RNA 干扰（RNAi）来实现染色质沉默。如下所述，裂殖酵母的研究已经揭示出这种沉默染色质对各特定异染色质区域（如着丝粒和接合型位点）的功能是必需的。在着丝粒，异染色质保证染色体的正常分离（Allshire et al. 1995；Ekwall et al. 1995）；在接合型位点，它有助于调节接合型的转换（Egel 2004 综述）。临近端粒处也有沉默染色质形成，虽然还未知其功能（Nimmo et al. 1994；Kanoh et al. 2005）。插入到 rDNA 中的标记基因也是沉默的（Thon and Verhein-Hansen 2000；Cam et al. 2005）。

与链孢霉和较高等的真核生物相比，裂殖酵母中没有检测到 DNA 甲基化（Wilkin-

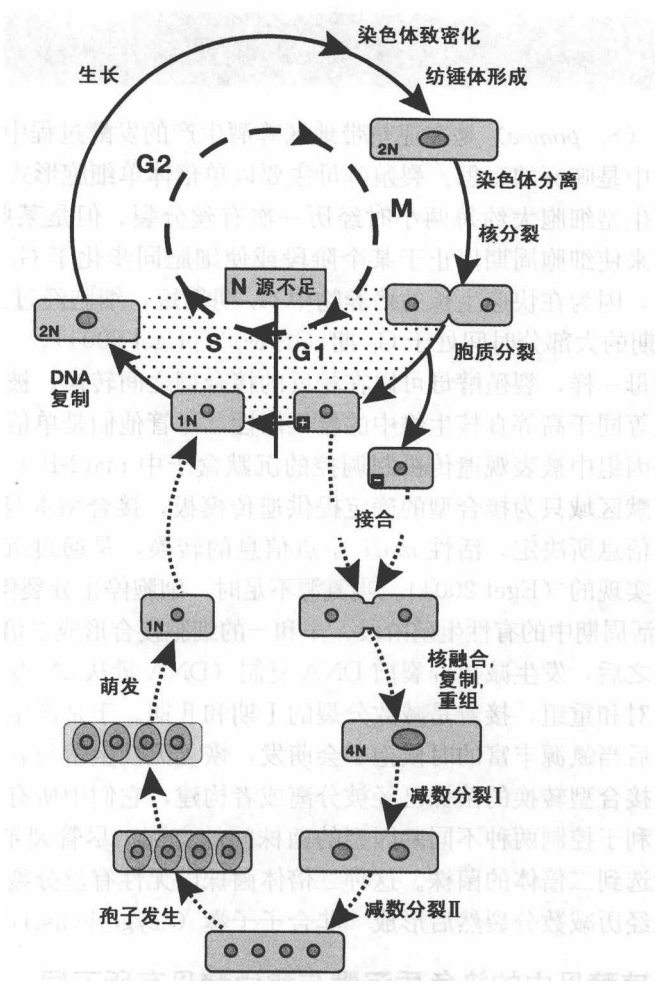



图 6-1 裂殖酵母的细胞周期

裂殖酵母的 $G_1$ 期很短，在细胞周期中少于 10%（为了利于描述，点状的区域在图中被扩大了）。在营养丰富的培养基中，$G_1$ 细胞进入 S 期，之后是一个长的 $G_2$ 期（约占细胞周期的 70%），随后是核分裂和细胞质分裂。当氮源不足时，相反接合型（+和−）的细胞结合，然后发生核融合。减数分裂前复制和同源重组使减数分裂Ⅰ期和Ⅱ期得以进行，产生四个单倍体的核，随后分别形成四个孢子并被包在子囊中。营养丰富时，孢子可以萌发，恢复细胞的无性生活周期。

son et al. 1995)，而这在很多真核生物中是染色质沉默的一种通用机制（见第 9 章和第 18 章)；因此在裂殖酵母中主要依赖染色质修饰来实现染色质沉默。如下面将要讨论的，这些修饰的建立和随后的染色质沉默利用了 RNAi 机制。

## 1.2 裂殖酵母着丝粒处的基因是沉默的

裂殖酵母着丝粒异染色质对核分裂时染色体的正常分离是必需的。研究揭示着丝粒实际上包含两个不同的染色质结构：异染色质和含有 CENP-A 的动粒染色质。通过研究插入到不同染色质区域的报告基因的不同沉默状态可以证实这一点。

在 DNA 水平，裂殖酵母着丝粒包括外侧重复序列（细分为 dg 和 dh，或者 K 和 L 元件）及其包围着的中央区域，中央区域包括内部重复序列（imr 或 B）和一个中央核心（cnt 或 CC）（图 6-2a）。染色体Ⅰ、Ⅱ和Ⅲ上的三个着丝粒 *cen1*、*cen2* 和 *cen3* 大约分别为 40kb、60kb 和 120kb（综述见 Egel 2004；Pidoux and Allshire 2004）。裂殖酵母着丝粒 DNA 的重复状态类似多细胞动物中更大更复杂的重复结构，但是其更易被操作（Takahashi et al. 1992；Steiner et al. 1993；Ngan and Clarke 1997）。在其他真核生物中，重复 DNA 通常和异染色质相关（另见第 5 章），所以在裂殖酵母着丝粒中的重复 DNA 意味着它们或许有异染色质的性质，如可以抑制基因表达。如下所述，每个裂殖酵母着丝粒的中央区域都由两段异染色质包围起来。而中央区域则装配成一种不同于邻近异染色质的染色质类型（CENP-A 染色质）。

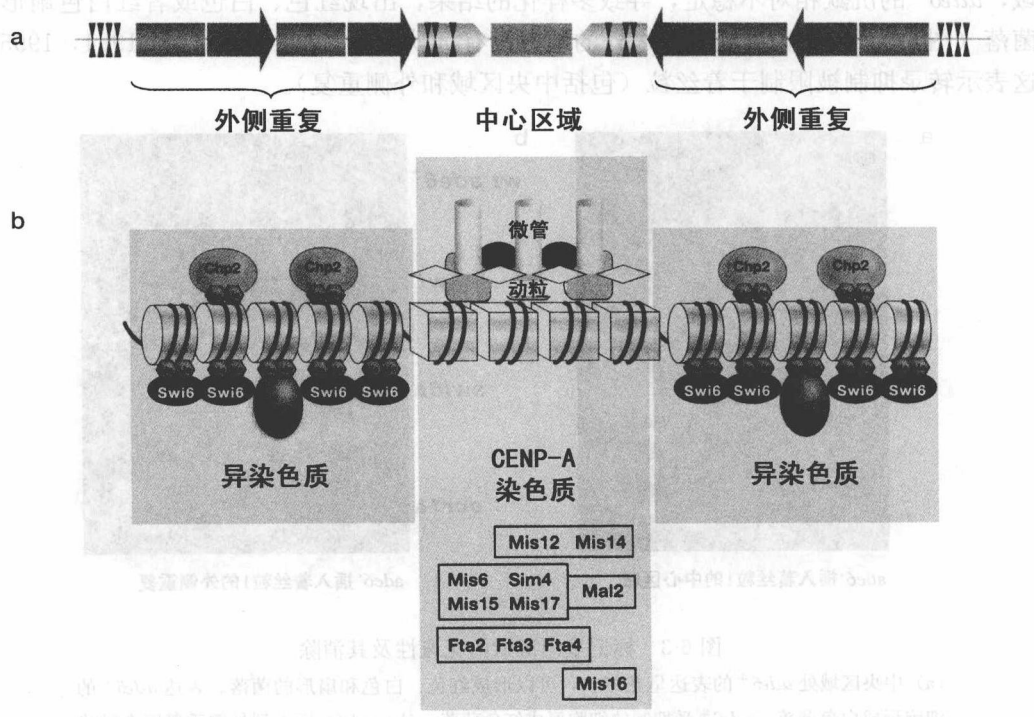



图 6-2　裂殖酵母着丝粒中不同的外侧重复异染色质和中央动粒结构

(a) 图中表示的是一个裂殖酵母的着丝粒。中央区域（粉色，动粒）包含 *imr* 和 *cnt* 元件，外侧重复包含转录的 dg 和 dh 重复（绿色，异染色质）。三个着丝粒有总体上相似的结构，但是外侧重复的数量上还是有所不同：*cen1*（40kb）有两个，*cen2*（65kb）有三个，*cen3*（110kb）有大约 13 个。tRNA 簇（双箭头表示）存在于 *imr* 区以及所有三个着丝粒的两侧。外侧重复和中央区域处的标记基因是转录沉默的。(b) 异染色质：外侧重复序列包装成的核小体上 H3K9 是被二甲基化的，从而为 chromo 结构域蛋白 Chp1、Chp2 和 Swi6 提供结合位点。中央动粒染色质：在中央区域 CENP-A 可能取代了 H3 形成了一种特殊的核小体（粉色方形）。除了 CENP-A，一些动粒蛋白（如图所示）也结合到中央区域序列而不与外侧重复结合。动粒在中央区域装配，其在纺锤体形成与染色体分离中介导与微管的结合。异染色质组分的突变可以减轻外侧重复处标记基因的沉默，但对中央区域没有影响。一些动粒组分的缺陷使中央区域处标记基因表达，却不影响外侧重复。

人们已经知道一个基因周围的染色质类型可以强烈的影响其表达，这最早在黑腹果蝇中被证明，染色体重排使“*white*”基因邻近着丝粒异染色质，导致其在复眼中不同的表达模式，结果眼睛表现出花斑状着色（图 5-1）。很明显，在许多生物体中，转基因要受到其插入位点周围环境的影响。

在裂殖酵母中，基因沉默可通过观察表型来检测，这和在酿酒酵母中评估报告基因表达的方法类似。例如，如 *ura4*$^+$报告基因被沉默，就可以形成抗 5-氟乳清酸的菌落。另外，*ade6*$^+$报告基因的沉默形成红色而不是白色的菌落（图 6-3a）。将正常情况下能表达的基因，如 *ura4*$^+$或者 *ade6*$^+$置于着丝粒（以外侧重复和中央结构元件所定义）处会导致转录沉默。外侧重复处会产生很强的沉默，所以大多数细胞都形成报告基因沉默状态被稳定维持的菌落（如 *ade6*$^+$作为报告基因，会形成红色菌落）。然而，在中央区域，*ade6*$^+$的沉默相对不稳定，导致多样化的结果，出现红色、白色或者红白色扇形的菌落（图 6-3a）。外侧重复 1kb 之外的地方没有沉默发生（Allshire et al. 1994，1995），这表示转录抑制被限制于着丝粒（包括中央区域和外侧重复）。

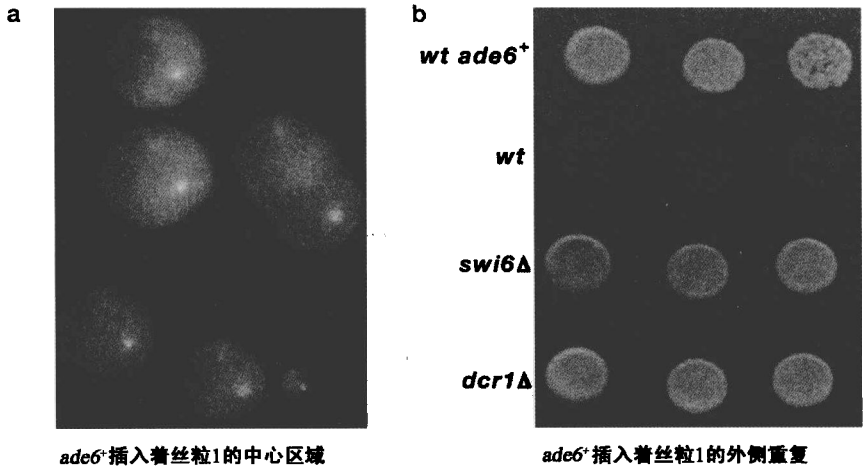



图 6-3 标记基因沉默的花斑性及其消除

(a) 中央区域处 *ade6*$^+$的表达呈花斑状，可以形成红色、白色和扇形的菌落。表达 *ade6*$^+$的细胞形成白色菌落，*ade6*$^+$受抑制的细胞形成红色菌落。(b) *ade6*$^+$插入到外侧重复区会被稳定沉默，形成均一的红色菌落。这种沉默需要异染色质蛋白，如 Swi6（结合甲基化的 H3K9），也需要如 Dcr1（Dicer，一种 RNaseⅢ型核酸酶）的 RNAi 组分。

## 1.3 裂殖酵母着丝粒由异染色质和中央动粒结构组成

裂殖酵母着丝粒中的不同沉默状态，是由外侧重复和中央区域中染色质结构以及与之结合的非组蛋白的差异所造成（Partridge et al. 2000）。在裂殖酵母着丝粒中发现两种不同的染色质结构：外侧重复区域的异染色质和中央区域的“CENP-A”染色质，动粒在后者装配。在这两个区域由不同的蛋白导致报告基因的沉默。

### 异染色质

染色质中组蛋白 H3 和 H4 的 N 端可以发生一系列的翻译后修饰，这通常与其处于活跃或抑制状态有相关性（见第 3 章）。外侧重复区域异染色质上具有组蛋白 H3 第九位赖氨酸的二甲基化及三甲基化（H3K9me2 和 H3K9me3）（Nakayama et al. 2001；Yamada et al. 2005）。着丝粒异染色质的形成需要几种蛋白使染色质发生修饰，这些修饰进而招募其他的因子。首先需要组蛋白去乙酰化酶（HDAC；如 Clr3、Clr6 和 Sir2）催化组蛋白 H3 的去乙酰化，以允许组蛋白赖氨酸甲基转移酶 Clr4 随后催化组蛋白 H3 第九位赖氨酸的甲基化。这种修饰为 Swi6 和 Chp2（第 5 章中描述的异染色质蛋白 1［HP1］的同源物）的 chromo 结构域以及另一种 chromo 结构域蛋白 Chp1 提供结合位点（图 6-2）。这些蛋白都有助于形成外侧重复区域沉默异染色质（Allshire et al. 1995；Cowieson et al. 2000；Partridge et al. 2000；Bannister et al. 2001；Nakayama er al. 2001；Shankaranarayana et al. 2003；Sadaie et al. 2004）。

插入到外侧重复区域的报告基因的染色质结构（如 *ura4*$^+$ 基因）显著富含 H3K9me2 和 Swi6 蛋白。这表明染色质修饰和 Swi6 可以从邻近的着丝粒重复 DNA 蔓延到插入的序列（Cowieson et al. 2000；Nakayama et al. 2001）。Swi6 的定位依赖于 H3K9 甲基化，在缺少 H3K9 专一性甲基转移酶 Clr4 或者 H3 第九位赖氨酸突变成精氨酸的细胞中 Swi6 的定位被破坏（Ekwall et al. 1996；Mellone et al. 2003）。Swi6 蛋白通过其 chromoshadow 结构域形成二聚体（Cowieson et al. 2000），这可能有助于它在 HDAC 和 Clr4 H3K9 甲基转移酶的帮助下沿着染色质纤维蔓延。

除 Swi6 之外，Chp1 和 Chp2 两种 chromo 结构域蛋白也通过与甲基化的 H3K9 相互作用，从而结合在着丝粒外侧重复区域染色质上。Chp1 被证明是 RNAi 复合物 RITS 的组分（见第 8 章），并且为完全建立外侧重复区域和被插入的报告基因处组蛋白的 H3K9 甲基化所必需（Partridge et al. 2002；Motamedi et al. 2004；Sadaie et al. 2004）。

### 中央动粒结构染色质

在讨论着丝粒外侧重复区域异染色质建立的细节之前，了解中央区域染色质是很重要的，此处和外侧重复区域有显著不同，因为动粒是在此处装配的。与外侧重复区的沉默有所不同，在 *cenl* 大约 10kb 的中央区域中，报告基因的沉默不依赖于 Clr4，所以也和组蛋白 H3K9 甲基化无关。事实上，中央区域有独特的染色质结构。这最开始被微球菌核酸酶分析所证明（见第 5 章对 MNase 消化的解释），分析结果显示为弥散的条带，而外部重复染色质显示出有规律的间隔 150bp 的梯状条带（Polizzi and Clarke 1991；Takahashi et al. 1992）。这种不同的形式有别于异染色质和常染色质，这和在此处装配的独特的 CENP-A 染色质以及动粒有关。在所有被检测的真核生物中，一个叫 CENP-A（或 cenH3）的类 H3 蛋白特异存在于有活性的着丝粒（cleveland et al. 2003）处，它在确定动粒的装配位点中起关键作用。

在裂殖酵母着丝粒中央区域，绝大部分组蛋白 H3 被 CENP-A 的直系同源物 Cnp1

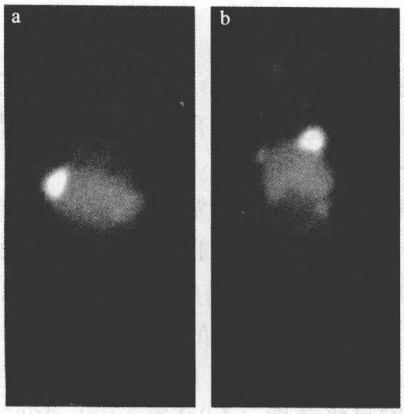



图 6-4 裂殖酵母核内的沉默染色质
两个间期的核，用荧光免疫定位显示异染色质（着丝粒、端粒和沉默的 *mat2-mat3* 位点）分布，红色为 Swi6，绿色为 $CENP^{Cnp1}$。不在绿色附近的红色信号代表着丝粒和 *mat2-mat3*。所有的着丝粒成簇分布在核周临近纺锤极体的位置。

所代替（图 6-4b）（Takahashi et al. 2000）。$CENP\text{-}A^{Cnp1}$ 可以在 S 期以复制依赖的方式装入染色质，也可以在 $G_2$ 期通过不依赖复制的方式被置入（详见第 13 章）。动粒蛋白本身可以调控 $CENP\text{-}A^{Cnp1}$ 的定位与装配（Goshima et al. 1999；Takahashi et al. 2000；Pidoux et al. 2003）。如 Ams2（一个 GATA 因子）在 S 期指导伴随复制的 $CENP\text{-}A^{Cnp1}$ 的组装。在缺乏 Ams2 的情况下，可以在 $G_2$ 期通过由 Mis6 指导的不依赖复制的方式来补偿，从而允许在间期完成 $CENP\text{-}A^{Cnp1}$ 的组装（Chen et al. 2003；Takahashi et al. 2005）。如果含 $CENP\text{-}A^{Cnp1}$ 的染色质结构被破坏（当 *cnp1*、*mal2*、*mis6*、*sim4* 和其他一些基因被突变），在微球菌核酸酶分析中弥散的结果转变成为更有代表性的结果（如核小体阶梯）。在影响中央区域染色质的突变体中并没有检测到外侧重复的沉默状态发生变化（Cowieson et al. 2000；Jin et al. 2002；Pidoux et al. 2003；Hayashi et al. 2004）。而且 $CENP\text{-}A^{Cnp1}$、*mal2*、*mis6*、*sim4* 和其他的动粒蛋白仅仅和中央区域染色质结合，这进一步证明了中央动粒区域在结构上更复杂，在功能上也和外侧重复染色质不同（图 6-2b）。

并非所有的动粒蛋白都被检测过，但是似乎中央区域的沉默是一个完整动粒装配的结果，像在酿酒酵母中一样，这至少涉及 50 个蛋白（Measday and Hieter 2004）。这巨大的蛋白复合体可能会限制 RNA 聚合酶Ⅱ接触到这个区域的报告基因从而阻止其转录。在许可温度下，一些突变体（如 *cnp1*、*mal2*、*mis6* 和 *sim4*）中动粒完整性会遭到部分破坏，这提高了报告基因的转录水平。这种现象的一个应用是可以用报告基因来检测中央核心结构的缺损，这导致了新的动粒蛋白被发现（Pidoux et al. 2003）。

奇怪的是，着丝粒区域并非完全没有基因，一些 tRNA 基因存在于外侧重复和中央动粒之间（图 6-2a）（Kuhn et al. 1991；Takahashi et al. 1991）。最近，这被证明可以作为屏障阻止异染色质侵入中央区域（Scott et al. 2006）。

## 1.4 着丝粒外侧重复序列本身导致了沉默染色质的装配

Clr4 依赖的沉默染色质装配不仅仅发生在着丝粒，也发生在约 20kb 的沉默接合型位点（*mat2-mat3*）（Noma et al. 2001）以及邻近端粒末端重复（由端粒酶合成的 TTACAGG）处（Nimmo et al. 1994，1998；Allshire et al. 1995；Kanoh et al. 2005）。存在于 *mat2* 和 *mat3* 之间的 CenH（centromere homologous，着丝粒同源）区域拥有一个和着丝粒外侧重复高度相似的大于 7kb 的序列（Grewal and Klar 1997）。另外，邻近染色体Ⅰ和Ⅱ端粒的 40kb 的端粒相关序列（telomere-associated sequences TAS）中有至少 0.5kb 的 DNA 序列和 CenH 有超过 84%的一致性。这说明着丝粒外侧

和相关的 CenH 重复可能以顺式作用的方式指导沉默染色质装配。

当着丝粒外侧重复（*dg*）或 *mat2-mat3*（CenH）的 DNA 序列被插入到正常情况下不发生沉默的基因组区域时，标记基因附近也可以发生 Clr4 依赖的沉默染色质装配（异位的沉默，图 6-5）（Ayoub et al. 2000；Partridge et al. 2002；Volpe et al. 2003）。一个简单的解释是，DNA 结合蛋白可以识别这个重复序列，当与之结合后这个蛋白招募 HDAC 和组蛋白赖氨酸甲基转移酶 Clr4，随后依次发生 H3K9 甲基化、chromo 结构域蛋白的结合和异染色质的形成。然而，真实情况要比这复杂。现在已经知道着丝粒外侧重复可以发生转录。令人瞩目的是，这种转录产生一个双链 RNA（dsRNA），其可以作为 RNAi 机器的底物，进而招募 Clr4 启动沉默染色质的组装（Volpe et al. 2002；Sadaie et al. 2004）。RNAi 同时也在相关的 CenH 重复和端粒的 TAS 重复处起作用（Cam et al. 2005；Kanoh et al. 2005）

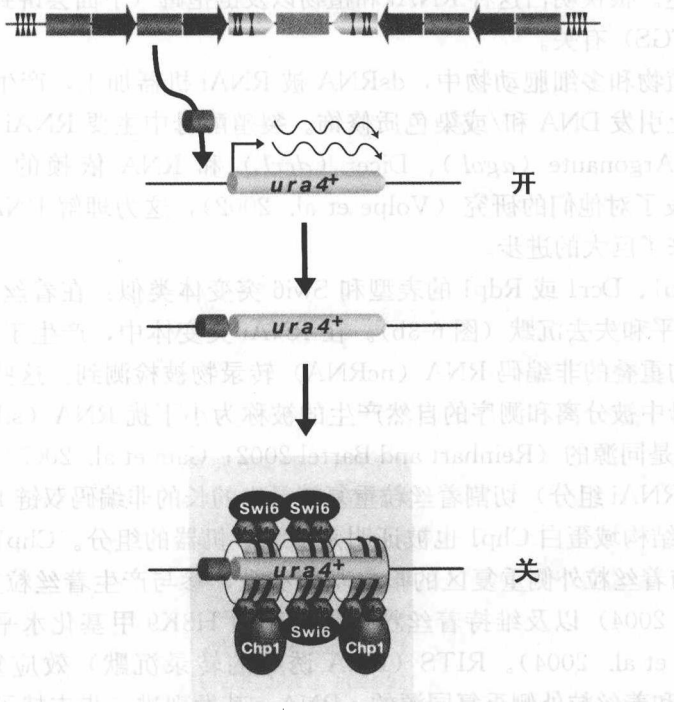



图 6-5　着丝粒重复序列介导沉默

在正常情况下不发生沉默的区域中具转录活性的 *ura4*$^+$ 基因周围插入 1～2kb 的重复序列片段，可导致 H3K9 甲基化、Swi6 和 Chp1 的结合并发生转录沉默。

然而，以下的发现使图景变得更加复杂：在 *mat2-mat3* 处 RNAi 发挥作用建立异染色质，然而在缺少核心 RNAi 组分时 DNA 结合蛋白 Atf1 和 Pcr1 结合在 CenH 附近来维持 *mat2-mat3* 处的沉默染色质（Jia et al. 2004；Kim et al. 2004）。类似的，在端粒处也发现冗余的沉默机制，末端重复自身可以通过结合在其上的 Taz1 招募 Clr4 进而招募 Swi6，但是 RNAi 也可以依赖 TAS 元件上的 CenH 部分来发挥作用，在端粒形成扩

大的沉默染色质区域（Nimmo et al. 1994；Allshire et al. 1995；Kanoh et al. 2005）。在着丝粒异染色质的维持中是否也有这种冗余机制呢？虽然在细胞缺少 RNAi 时着丝粒外侧重复不能转录自身，报告基因也不再沉默，但是 H3K9me2 依旧存在（Sadaie et al. 2004）。是什么因子维持这种甲基化？一种可能是 CENP-B 相关蛋白（Abp1、Cbh1、Cbh2），他们有可以结合到着丝粒重复 DNA 的保守的 DNA 结合结构域，并为沉默染色质的有效形成所必需（Irelan et al. 2001；Nakagawa et al. 2002）。其他的发现表明 Clr3 HDAC 在维持异染色质完整性中的作用也不依赖于 RNAi（Yamada et al. 2005）。

## 1.5 RNA 干扰指导沉默染色质装配

RNAi 现象首次发现于秀丽隐杆线虫，当表达双链 RNA（dsRNA）时会导致同源基因失去表达。很快明白这种 RNAi 和植物以及链孢霉（下面会讲到）描述的基因转录水平沉默（TGS）有关。

已知在植物和多细胞动物中，dsRNA 被 RNAi 机器加工，产生的小 RNA 可以在同源染色质上引发 DNA 和/或染色质修饰。裂殖酵母中主要 RNAi 组分直系同源基因的存在，如 Argonaute（*agol*）、Dicer（*dcrl*）和 RNA 依赖的 RNA 聚合酶基因（*rdpl*），激发了对他们的研究（Volpe et al. 2002），这为理解 RNAi 介导的染色质修饰和沉默带来了巨大的进步。

失去 Agol、Dcrl 或 Rdpl 的表型和 Swi6 突变体类似：在着丝粒外侧重复区降低 H3K9me2 水平和失去沉默（图 6-3b）。在 RNAi 突变体中，产生于着丝粒外侧重复的不连续大小的重叠的非编码 RNA（ncRNA）转录物被检测到。这些非编码 RNA 和那些从裂殖酵母中被分离和测序的自然产生的被称为小干扰 RNA（siRNA，约 21nt）的小 RNA 片段是同源的（Reinhart and Bartel 2002；Cam et al. 2005）。这些 siRNA 是由 Dicer（一种 RNAi 组分）切割着丝粒重复转录出的长的非编码双链 RNA 所产生的。

Chromo 结构域蛋白 Chpl 也被证明是 RNAi 机器的组分。Chpl 结合到 H3K9me2 染色质，参与着丝粒外侧重复区的报告基因沉默，参与产生着丝粒重复同源的 siRNA（Noma et al. 2004）以及维持着丝粒重复的正常 H3K9 甲基化水平（Partridge et al. 2002；Sadaie et al. 2004）。RITS（RNA 诱导的转录沉默）效应复合体包含 Chpl、Agol、Tas3 和着丝粒外侧重复同源的 siRNA，其发现进一步支持了 Chpl 在 RNAi 中的作用（Motamedi et al. 2004）。另一个含有 Rdpl 的蛋白复合体（RDRC：依赖 RNA 的 RNA 聚合酶复合体）也被确认，它含有一个预测的 RNA 解旋酶（Hrrl）和一个假定的多聚腺嘌呤核糖核苷酸聚合酶（Cid12）。这也似乎意味着完整的外侧重复区沉默染色质组装和 RNAi 是一个整体过程（图 6-6）（Motamedi et al. 2004）。

用 HDAC 抑制剂曲古抑菌素 A（trichostatin A，TSA）处理 RNAi 缺陷的细胞，发现了一种在沉默接合型位点维持异染色质的不依赖于 RNAi（依赖于 Atf1/Pcr1）的沉默机制。当在 RNAi 突变体中模仿缺失 *atf1* 或 *pcr1* 的效果时，*mat2-mat3* 区域完全失去了沉默染色质（Hall et al. 2002；Jia et al. 2004）。如前所述，着丝粒外侧重复区沉默染色质的形成不需要 Atf1 和 Pcr1（Kim er al. 2004）。TSA 的短暂处理也会导致

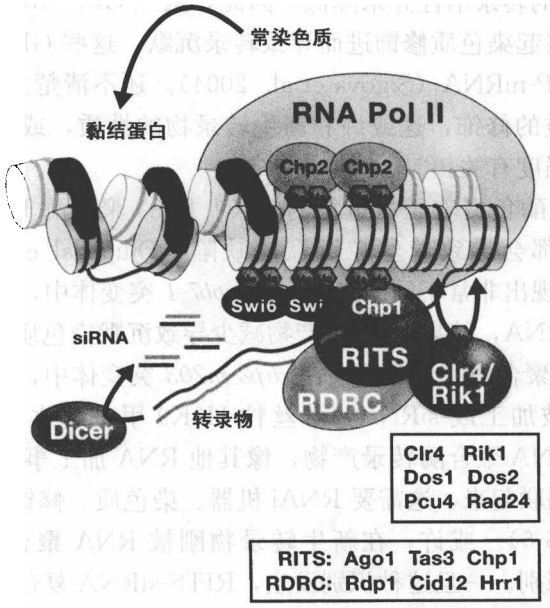



图 6-6　着丝粒重复处的转录把 RNAi、异染色质形成和染色体黏联联系起来

RNA 聚合酶Ⅱ催化的外侧重复的转录为 RNAi 和 Dicer 依赖的 siRNA 的产生提供最初的底物。RITS 复合体（Ago1、Tas3、Chp1）中的 Ago1 与 siRNA 的结合允许其靶向同源的转录物。RDRC（Rdp1、Cid12 和 Hrr1）合成的 dsRNA 为 Dcr1 提供了更多的底物以产生 siRNA，这或许有放大信号的效果。转录物、RNA 聚合酶Ⅱ亚基和 RNAi 组分间的相互作用招募了 Clr4，它是一个 H3K9 的甲基化梅，它可以导致 chromo 结构域蛋白和黏结蛋白的结合以及姐妹着丝粒间的黏联。

着丝粒沉默的丢失，以及高乙酰化的组蛋白和着丝粒功能的缺陷（Ekwall et al. 1997）。虽然这种 TSA 诱导的“表观状态”仅仅相对稳定，但是也能维持几轮细胞分裂甚至减数分裂。对此最可能的解释是，一旦这种不正常的高乙酰化状态发生，异染色质重新建立是困难的。表观状态也能在 *mat2-mat3* 位点被建立，同样也能在减数分裂中被维持（Nakayama er al. 2000）。

## 1.6　RNA 聚合酶Ⅱ合成的着丝粒重复转录产物使 RNAi 和染色质修饰相联系

前面已经说明了着丝粒异染色质的装配需要 RNAi 和组蛋白修饰，这种异染色质可以蔓延到邻近插入的报告基因。这就提出了一些显而易见的问题，例如，RNAi 如何与组蛋白修饰一起来导致沉默染色质的形成？这种 RNAi 指导的染色质修饰与沉默又如何靶向特定的基因组区域。

在很多生物体中，与某个基因同源的 dsRNA 的转录，会导致基因转录沉默（DNA/染色质修饰，TGS），或者转录后基因沉默（PTGS，mRNA 降解）。在裂殖酵母中任何 dsRNA 都可以形成 siRNA 吗，是否某种 siRNA 只能诱导转录后基因沉默，或者它们同时也可以引起染色质修饰（如 H3K9me2）进而导致同源基因的转录水平沉默？表达 GFP 报告基因同源的 dsRNA 产生的 siRNA 可以导致 GFP 转录产物的减少，

但是 GFP 报告基因的转录活性并未降低。因此，虽然 GFP-siRNA 减少了 GFP mRNA 的水平，但是不能引起染色质修饰进而导致转录沉默。这些 GFP-siRNA 一定是仅仅在转录后水平破坏 GFP-mRNA（Sigova et al. 2004）。还不清楚为什么 GFP-siRNA 不能诱导同源基因染色质的修饰，这或许和新生转录物的性质，或者和 GFP 基因的 RNA 聚合酶Ⅱ启动子的强度有关。

何种 RNA 聚合酶负责着丝粒重复的转录？RNA 聚合酶 II 亚基（Rpb2 和 Rpb7）任何一个发生突变都会导致着丝粒沉默的缺陷（Djupedal et al. 2005；Kato et al. 2005），虽然它们表现出非常不同的表型。在 *rpb7-1* 突变体中，有较低的着丝粒重复转录水平，较少的 ncRNA，结果 siRNA 产物减少导致沉默染色质丢失。这表明着丝粒重复的转录需要 RNA 聚合酶Ⅱ。相反，在 *rbp2-m203* 突变体中，着丝粒转录产物正常产生，但是他们不能被加工成 siRNA，着丝粒 H3K9 甲基化水平降低。这些研究说明 RNAi 不仅仅需要 RNA 聚合酶转录产物，像其他 RNA 加工事件一样，着丝粒 siRNA 的产生必须同转录偶联起来，这需要 RNAi 机器、染色质、修饰酶和 RNA 聚合酶Ⅱ之间的相互作用（图 6-6）。或许，在新生转录物刚被 RNA 聚合酶Ⅱ合成的时候就被 RITS 中的 siRNA 识别。一旦这种识别发生，RITS-siRNA 复合体就牢固地结合在转录产物上，接下来可以招募染色质修饰相关的蛋白，如 Clr4。令人惊讶的是，在缺少 Clr4 组蛋白赖氨酸甲基转移酶活性的细胞中，着丝粒 siRNA 也被丢失（Noma et al. 2004；Hong et al. 2005）。或许 Clr4 的缺失可以削弱转录、RNAi 和染色质修饰的不同组分间结合的稳定性，从而影响了 siRNA 的产生（Mtamedi et al. 2004）。另一种可能的解释是，H3K9 甲基化以顺式作用的方式参与着丝粒初始转录物的各种 RNA 加工过程，从而影响 siRNA 的产生（图 6-6）（Noma et al. 2004；Sugiyama et al. 2005；见第 8 章）。

## 1.7 着丝粒沉默染色质参与姐妹染色体着丝粒黏联和染色体正常分离

着丝粒外侧重复异染色质和 CENP-$A^{Cnp1}$ 动粒染色质是如何影响染色体分离呢？Clr4 依赖的沉默异染色质装配，发生在着丝粒外侧重复以及与之相关的接合型位点和端粒附近的重续序列。用裸露的 DNA 质粒进行的实验表明通过某种方式外部重复在有功能的着丝粒的装配中起作用，可以使这些 Cen 质粒拥有在有丝分裂和减数分裂纺锤体中分离的能力。但是仅仅外部重复或者仅仅中央区域都不足以装配成一个有功能的着丝粒（Clarke and Baum 1990；Takahashi et al. 1992；Baum et al. 1994；Ngan and Clarke 1997；Pidoux and Allshire 2004）。

导致外侧重复区沉默丢失的突变（如缺少 Clr4、RNAi 组分或 Swi6）也会提高有丝分裂时染色体丢失的速度以及后期染色体滞后发生的频率（图 6-7）（Allshire et al. 1995；Ekwall et al. 1996；Bernard et al. 2001；Nonaka et al. 2002；Hall et al. 2003；Volpe et al. 2003）。缺失 Swi6 的细胞存在着丝粒黏联的缺陷但是仍然保持着染色体臂的黏联（Bernard et al. 2001；Nonaka et al. 2002）。合适的纺锤体的形成需要姐妹动粒结合到从纺锤体两极发散出的微管纤维上。作用在动粒上的拉力需要姐妹动粒紧紧结合（图 6-7b）。Siw6 参与招募黏结蛋白（cohesin）到外侧重复染色质，进而介导姐妹着丝

粒间的紧密结合（图 6-6）。因此着丝粒沉默染色质的一项功能是介导着丝粒间的黏联。

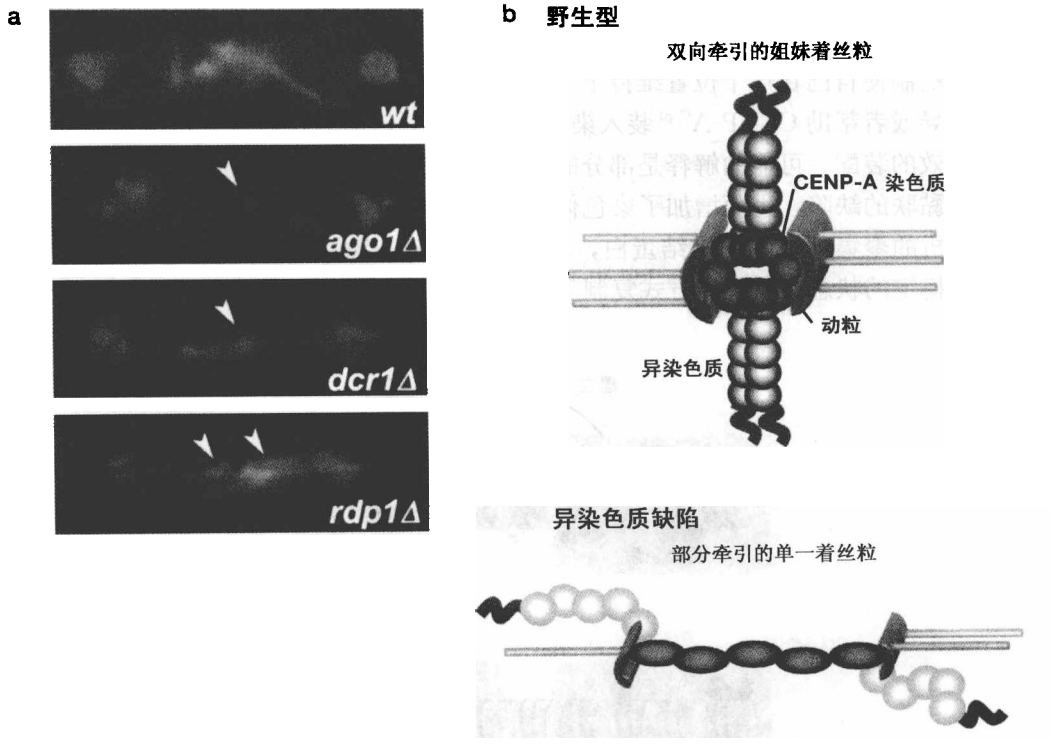



图 6-7　异染色质丢失导致染色体分离的缺陷

(a) 缺少 RNAi 和异染色质组分的细胞会加速染色体丢失并提高后期染色体滞后发生的频率。(b) 异染色质缺陷细胞导致染色体滞后可能是因为异常装配的动粒会同时与两极发出的微管相结合。这种方向性的不正确会一直持续到分裂后期，直到与某一极微管结合被破坏，引起随机分离，导致染色体的丢失或获得。

黏结蛋白与端粒和 *mat2-mat3* 位点也有很强的结合（Bernard et al. 2001；Nonaka et al. 2002）。另外，黏结蛋白也结合到邻近异位着丝粒重复的 *ura4*$^+$ 基因处的沉默染色质（图 6-5），这表明了沉默染色质和黏结蛋白之间的联系（Partridge et al. 2002）。因此，黏结蛋白的招募似乎是 Swi6 结合的异染色质的一个普遍特性。还不清楚 Swi6 染色质如何招募黏结蛋白，但是 Swi6 确实可以和黏结蛋白亚基 Psc3 相互作用（Nonaka et al. 2002）。保守的激酶 Hsk1（Cdc7）的调节亚基 Dfp1 可以和 Swi6 相互作用，并且是招募黏结蛋白到着丝粒所必需的（Bailis et al. 2003）。既然在脊椎动物细胞中缺失 RNAi 组分 Dicer 会影响异染色质完整性和姐妹着丝粒间的结合，那么说明异染色质和染色单体黏联之间的关系在其他生物体中也是保守的（Fukagawa et al. 2004）。

## 1.8　功能性着丝粒表观遗传学状态的继承

通过研究含有最少着丝粒区域的质粒如何装配成有功能的着丝粒，发现了一个有趣的表观遗传现象。虽然仅含有一部分外侧重复和大多数中央区域的质粒装配成有功能的着丝粒的效率很低，但是奇怪的是，一旦这种有功能的着丝粒状态被建立，它就能在多

次有丝分裂甚至减数分裂中保持下来（Steiner and Clarke 1994；Ngan and Clarke 1997）。一个解释是外侧重复为动粒装配提供了一个有利环境（Pidoux and Allshire 2005），但是一旦装配好了，CENP-A$^{Cnp1}$染色质及动粒结构自身就会通过一种和复制相关的模板机制使自己在这个位置维持下来（Tkahashi et al. 2005）。或许异染色质以某种方式诱导或者帮助 CENP-A$^{Cnp1}$装入染色质（图 6-8），但是仅有部分的异染色质不能使动粒高效的装配。可能的解释是部分的外侧重复不足以招募足够的黏结蛋白，这导致着丝粒间黏联的缺陷，进而增加了染色体丢失的比率。这种着丝粒结构可能在几次细胞分裂中随机的聚集到足够的黏结蛋白，增加了在有丝分裂中的稳定性。一旦这些被完成，这种稳定的状态会以某种方式复制到子代分子，使其能在细胞分裂中得到维持（图 6-8）。

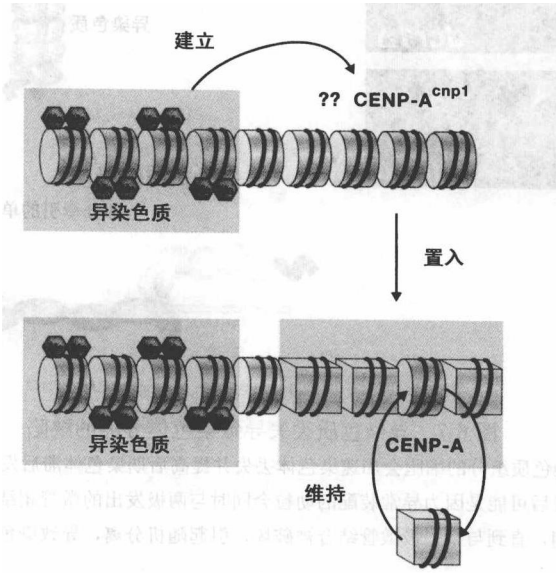



图 6-8 CENP-A 染色质的建立和维持

中央区域的 DNA 自身并不能装配功能性的着丝粒，而需要外侧重复。已经装配好的着丝粒失去异染色质并不影响中央区域 CENP-A$^{cnp1}$或动粒的结构。这表明异染色质或许以某种方式指引 CENP-A$^{cnp1}$染色质以及动粒组装的定位。CENP-A$^{cnp1}$如何装入核小体或者这种特殊的染色质结构在中央区域如何维持还依然未知。

如前所述，TSA 可以引起沉默丢失从而导致着丝粒缺陷（Ekwall et al. 1997）。考虑到沉默染色质形成与着丝粒黏联间的关系，似乎 TSA 引起的高乙酰化阻止了依赖 RNAi 的沉默染色质的高效重建，因为丢失了沉默的状态以及姐妹着丝粒黏联的有缺陷的着丝粒被传到了子代（详见第 14 章）。

## 1.9 真菌中沉默机制的多样性

本章前半部分，我们描述了在一种相对简单的单细胞真核生物裂殖酵母中，报告基因如何被着丝粒上两种不同类型的染色质所沉默。CENP-A 染色质存在于着丝粒中央，两边被异染色质包围。CENP-A 染色质标记着动粒组装的位置，推测此处与微管连接。

在两边异染色质形成中，RNAi 识别从外侧重复转录出的非编码 RNA，招募来染色质修饰酶，如 Clr4，从而产生了 chromo 结构域蛋白的结合位点，然后导致强烈的沉默。这种异染色质为着丝粒提供了紧密的黏联，也帮助动粒装配。RNAi 在着丝粒沉默染色质形成中的作用起源于其在基因组抵御 RNA 病毒和转座元件方面的作用，就如在植物中发现的那样。事实上，着丝粒重复可能是古老可移动元件的遗迹。

这种推测被丝状真菌链孢霉（本章下半部分的主题）的早期研究所支持。虽然在不同的真菌中很明显有共同的表观遗传机制，但是如下所述，也存在很多差异。例如，不同于已被很好研究的芽殖和裂殖酵母，链孢霉有 DNA 甲基化，这在高等真核生物中是一种经典通用的表观遗传标记，如在哺乳动物和开花植物中。此外，链孢霉的研究揭示出几种在生命周期不同阶段起作用的独立的沉默系统。第一种叫重复序列诱导的点突变（RIP），它同时涉及表观遗传学机制和遗传学机制，作为一种基因组防护机制起作用。第二种叫平抑（quelling），是一种以 RNAi 为基础的机制，可以导致转基因及其内源同源基因的沉默。第三种叫非配对 DNA 导致的减数分裂沉默（MSUD），也是以 RNAi 为基础，但在起作用的时间、靶点和目的上与平抑不同。虽然在包括本章讲述的模式真菌在内的所有生物体中，我们对表观遗传学的研究尚处于初始阶段，但是对裂殖酵母和链孢霉的研究一定会为理解真核生物的表观遗传机制提供丰富的信息。

## 2. 链孢霉：历史和特性

20 世纪 20 年代晚期，Dodge 第一次把丝状真菌链孢霉（图 6-9 和 6-10）当作实验材料。大约十年后，Beadle 和 Tatum 用它进行了著名的“一个基因一个酶”的研究，这把生物化学和遗传学联系起来（Davis 2000）。Beadle 和 Tatum 选择其进行研究的一部分原因是因为它生长速度快、易于在成分确定的培养基中繁殖以及易于进行遗传操作（如突变、互补测验和基因定位）。尽管不像某些真核模式生物那样被广泛研究，但是链孢霉还是继续吸引着一些研究人员，因为它的适度的复杂性，也因为它很适合进行一些遗传、发育和亚细胞方面的研究（Borkovich et al. 2004）。它尤其适合研究光生物学、生物节律、种群生物学、形态建成、线粒体侵入、DNA 修复和重组、DNA 甲基化和其他的表观遗传过程。

图 6-9 描述了链孢霉的生活周期。无性生殖阶段起始于一个有性孢子（子囊孢子）或者一个无性孢子（分生孢子）的萌发，接着形成多核细胞组成的丝状分支（菌丝；图 6-10C）。两种接合型（A 和 a）在形态学上不可区分（图 6-10B）。子囊孢子的萌发需要热激，分生孢子可以自然萌发。菌丝生长很快（37℃生长速度大于每小时 5mm），最后会形成菌丝体。菌丝体形成后，空中的菌丝进一步生长，导致形成大量橙色的分生孢子，这是此物种的标志性特征（图 6-10A、B）。这些含有一个或多个核的分生孢子，或者形成新的无性个体，或者使异性受精。当受营养条件限制时，链孢霉通过形成成熟的子实体（“原子囊”）而进入有性生殖阶段。当从原子囊生出的特化的菌丝（“受精丝”）接触到相反接合型的组织时，会接受一个核并将其送回原子囊。链孢霉及其他的丝状子囊菌在有性生殖阶段与酵母不同，他们在受精和核融合之间有一个很长的异型核阶段。

异型核细胞在子囊壳中增殖。在最后的分裂时，细胞是双核的，两种接合型各一个。这些细胞会形成钩状（“拐杖”），最后的有丝分裂产生四个核，然后隔膜形成，在拐杖的弯曲部分出现一个双核细胞。这个细胞会形成一个子囊。遗传分析显示，一般情况下，一个子囊壳中会产生约 100 个或者更多的子囊，它们来源于共同的父本和母本。当核融合之后，形成的二倍体核立即进入减数分裂。因此，生活周期中的二倍体阶段是短而有限的。减数分裂产物在被包裹作为子囊孢子之前会进行一次有丝分裂，后来在子囊孢子发育时中还会进行有丝分裂（图 6-9，Davis 2000）。

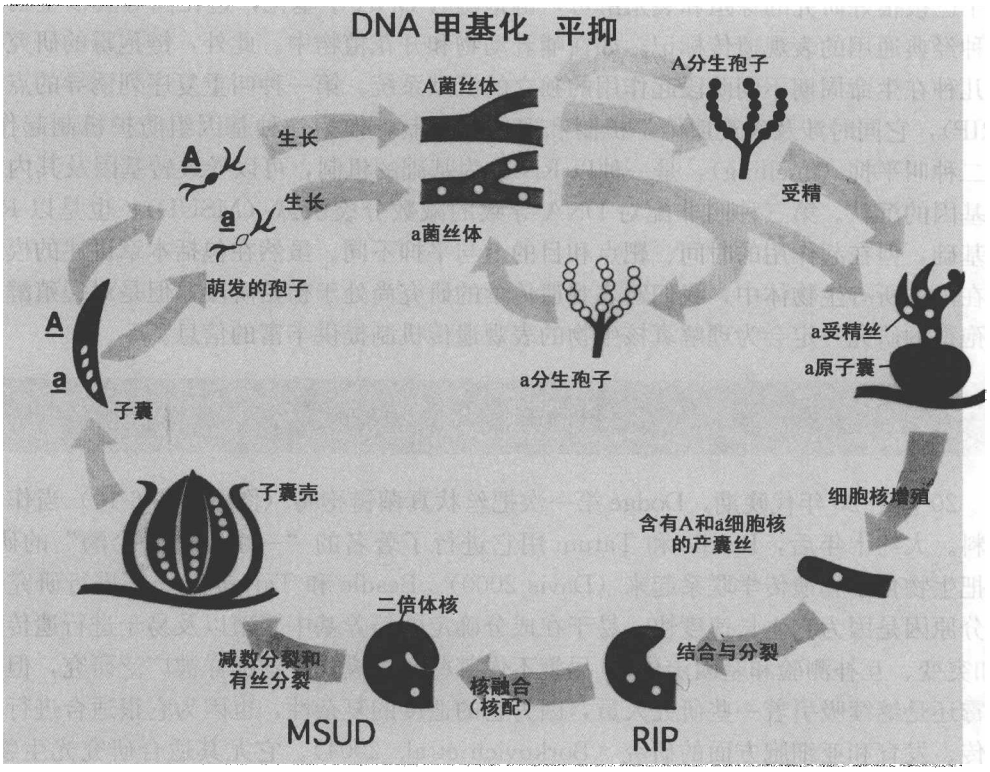



图 6-9 链孢霉的生活周期

文中所述的各表观遗传过程发生的阶段见本图（经许可，重印自 Shiu et al. 2001）。

链孢霉的基因组大约 40Mb，有 7 条染色体，预测有约 1 万个编码蛋白的基因（Galagan et al. 2003），粗略计算总共有 1000 个单位的遗传图谱长度（Perkins et al. 2001）。仅仅 10%的基因组由重复 DNA 组成，除了 70 个拷贝的编码三个大 rRNA 的 9kb 左右 rDNA 单元的串联阵列，大多数重复 DNA 含有失活的可移动元件。大多数链孢霉株缺少有活性的转座子，也极少有序列相近的旁系同源基因，这几乎理所当然的反映了 RIP 的作用，RIP 是在真核生物中发现的首个同源序列依赖的基因组防卫系统（Selker 1990）。我们知道链孢霉有至少三种用于保全基因组的基因沉默过程：RIP、平抑（quelling）和 MSUD（Borkovich et al. 2004）。所有这些过程都涉及表观遗传学机制也都直接或间接的和 DNA 甲基化（一个在链孢霉和很多真核生物中发现的基本的表

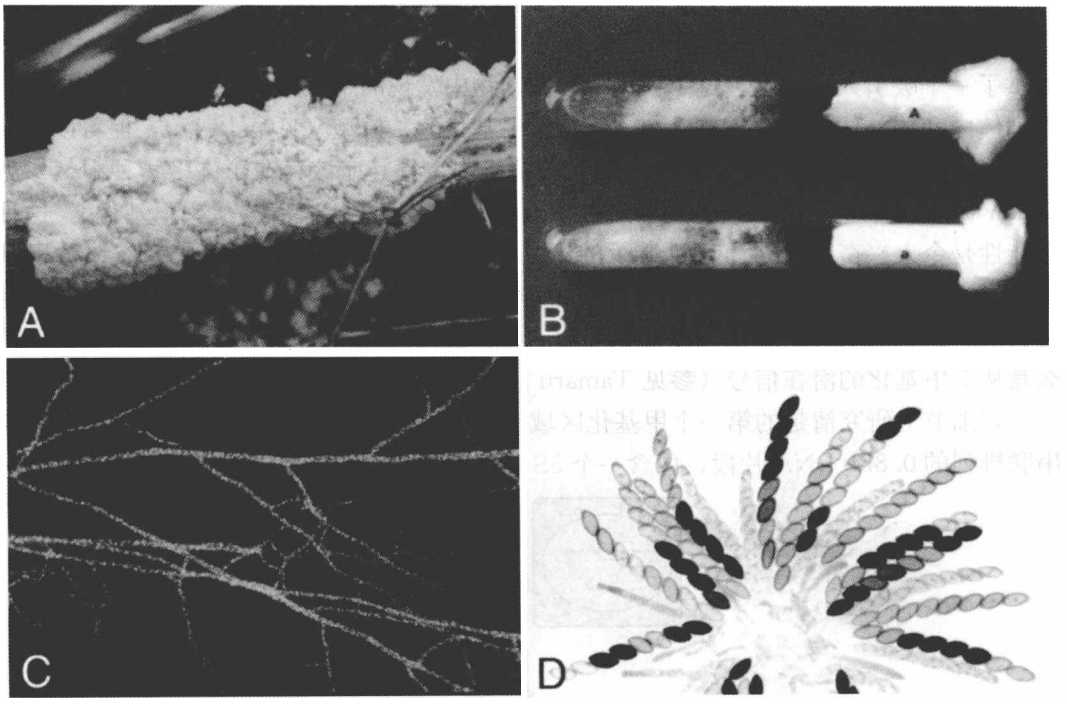



图 6-10　链孢霉的图片

(A) 野生于甘蔗上的营养生长的菌（来自斯坦福大学 D. Jacobson）。(B) 实验室斜面培养基上的营养生长的菌（来自斯坦福大学 N. B. Raju）。(C) 用 DAP1 染出大量细胞核的链孢霉菌丝（来自斯坦福大学 M. Springer）。(D) 成熟子囊形成的圆花窗图案（经许可，重印自 Raju 1980 [©Elsevier]）。

观遗传学机制）有关。我们将先讨论 DNA 甲基化，随后讨论 RIP、平抑和 MSUD。

## 2.1　链孢霉中的 DNA 甲基化

自从数十年前被发现以来，真核生物的 DNA 甲基化一直是个引人注目的谜。一些基本的问题依然存在争议，例如，是什么决定哪段染色质区域被甲基化？DNA 甲基化的功能又是什么？链孢霉为研究 DNA 甲基化的调控与功能提供了一个出色的模型。一些其他的真核模式生物，包括线虫类的秀丽隐杆线虫、酵母类的酿酒酵母和裂殖酵母都缺少 DNA 甲基化，在另一个模式生物黑腹果蝇中 DNA 甲基化的报道仍存在争议。在一些生物体中，DNA 甲基化为生存所必需，如哺乳动物，这使某些分析变得复杂。在链孢霉中，约 1.5%的胞嘧啶是被甲基化的，但是这种甲基化不是必需的，这为遗传学研究提供了便利。虽然当从一个系统推测到另一个系统时，我们必须小心谨慎，但是至少从某些方面来看 DNA 甲基化在物种间还是保守的。例如，所有已知的 DNA 甲基转移酶，包括在原核生物和真核生物中发现的，在他们的催化结构域都表现出很强的同源性（Grace Goll and Bestor 2005）。过去十年中，研究发现 DNA 甲基化的调控和功能在链孢霉、拟南芥、小鼠和其他系统中既有很重要的相似性也有有趣的不同点，这证明了在不同模式生物中进行研究的重要性。

链孢霉中 DNA 甲基化的发现最初令人感兴趣是因为它不局限于对称的位点，如 CpG 二核苷酸或 CpNpG 三核苷酸。Riggs、Holliday 和 Pugh 已经为甲基化继承和维持提供了一个吸引人的模型，这个模型依赖在动物中发现的甲基化位点的对称性。虽然一系列体内和体外的研究都支持这个“维持性甲基化酶”模型（见第 18 章），但可以想象存在着不依赖于对称位点的 DNA 甲基化维持机制，且其应该存在于许多物种中（参见 Selker et al. 2002）。被发现的在非对称位点的甲基化可能代表着“从头甲基化”，这种可能性是令人兴奋的，因为盲目继承的甲基化机制无法解释最初哪些序列应该被甲基化。事实上，在链孢霉中采用 DNA 介导的转化和甲基化抑制剂的研究结果证明了存在可被重新建立的从头甲基化（参见 Singer et al. 1995）。其他的研究至少部分确认了什么是从头甲基化的潜在信号（参见 Tamaru and Selker 2003）。

在细节上研究清楚的第一个甲基化区域是 1.6kb 的 ζ-η 区域，它包括两段有差异的串联排列的 0.8kb DNA 片段，内含一个 5S rRNA 基因。与缺少该重复序列的菌株的比

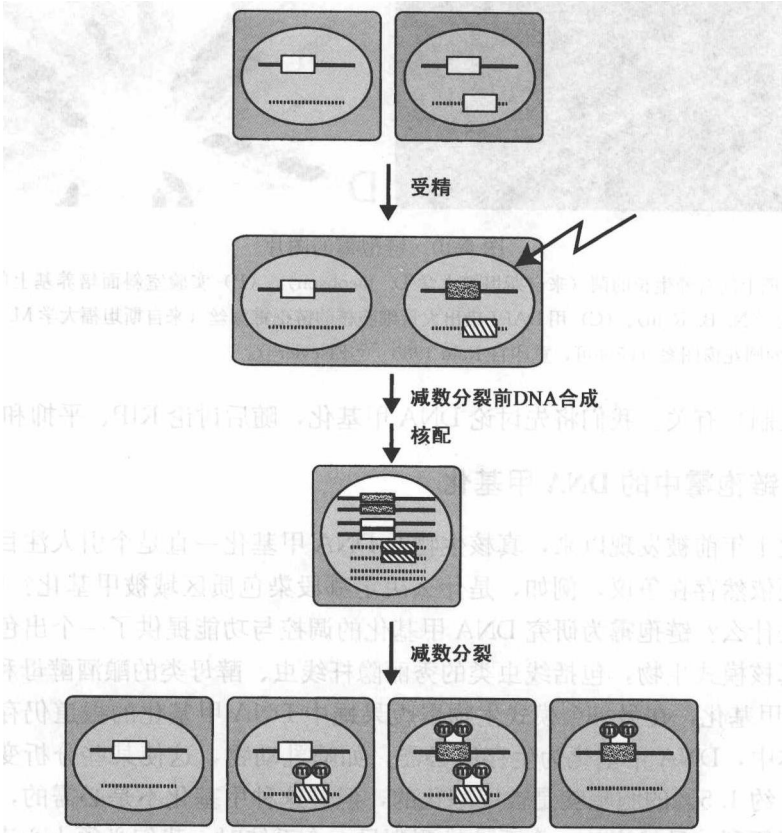



图 6-11 重复序列诱导的点突变（RIP）

为了清楚起见，只列出了两条染色体。空框代表一个基因或者任一染色体片段，当存在与其重复的序列时（如右上显示的菌株），其会在受精和核融合之间发生 RIP（用闪电图形表示）。遗传实验的结果显示重复序列可以在约 10 次有丝分裂期间不断地发生一系列 C 到 T 的转换（用填充的框表示），直到最后的前有丝分裂 DNA 合成（Selker et al. 1987；Watters et al. 1999）。底图所示为子代染色体的四种可能组合，红色“m”代表 DNA 甲基化，其与 RIP 产物高度相关（尽管不总是如此）。

较，导致提出了重复序列能通过某种方式诱导 DNA 甲基化的观点，最后发现了叫做 RIP 的基因组防御系统。RIP 的阐明揭示了重复序列并不直接启动 DNA 甲基化，至少在链孢霉中是这样，它启动的是 RIP（图 6-11），RIP 和 DNA 甲基化紧密相关。ζ-η 区域和 φ63 区域（链孢霉中第二个被发现的甲基化区域）都是 RIP 的产物。随后的全基因组甲基化分析更揭示链孢霉中大部分甲基化区域都是被 RIP 失活的转座子的遗迹（Galagan er al. 2003；Selker et al. 2003）。事实上，在链孢霉中唯一可能不经 RIP 介导的 DNA 甲基化存在于串联排列的 rDNA（Perkins et al. 1986）。

## 2.2　RIP——既涉及遗传学机制也涉及表观遗传学机制的基因组防卫系统

对链孢霉转化株杂交子代的精细分析导致发现了 RIP（Selker 1990）。研究发现，重复序列，不论是天然的还是外来的，连锁的或者非连锁的，都会在受精后产生的二核细胞阶段发生许多碱基转换突变（G：C 变成 A：T）（图 6-12）。对某基因单独存在于基因组或其与一非连锁的同源基因共存于基因组这两种情况下比较其稳定性，结果显示 RIP 不仅仅与重复相关，事实上它正是由重复所诱导的。经历一次有性生殖周期，重复基因中多至 30%的 G：C 碱基对可以被突变。被 RIP 改变的序列高频率地（但并非毫无例外地）发生从头 DNA 甲基化。似乎源于 RIP 的突变是由 5-甲基胞嘧啶或者胞嘧啶发生了酶促脱氨反应，而后进行复制所导致的（Selker 1990）。胞嘧啶甲基化会经历一个易于发生自发性脱氨基反应的中间步骤，这提示了或许 RIP 中可能存在的脱氨基反应是由 DMT 或者类 DMT 的酶来催化的。与这种可能性相符，从链孢霉基因组序列预测出的两个 DMT 同源基因中有一个与 RIP 有关（Freitag et al. 2002）。*Rid*（RIP 缺陷）基因缺陷纯合子间杂交产生的子代，不再发生新的 RIP。*Rid* 突变体并没有在 DNA 甲基化、受精、生长和发育中表现出明显的缺陷。相反，第二个链孢霉 DMT 同源基因（DIM-2）通过遗传学方法被确认为所有已知的 DNA 甲基化所必需，但并不被 RIP 所需要（Kouzminova and Selker 2001）。

所有的迹象都表明每个合适大小的重复（大于 400bp 的相连重复或者大于 1000bp 的非相连重复）都会在特殊的双核细胞期的某阶段经历 RIP。不过，重复会以一定几率逃脱 RIP（一般相连重复小于 1%，非相连重复小于 50%）。即便含有许多基因的染色体片段的重复也对 RIP 敏感（Perkins et al. 1997；Bhat and Kasbekar 2001）。虽然 RIP 被限制在有性生殖阶段，这个过程的存在还是提出了一个问题，是否链孢霉可以利用基因倍增来进化。基因组序列揭示了基因家族的存在，但几乎所有的旁系同源基因都产生了足够的分歧，以致不能启动 RIP（Galagan and Selker 2004）。我们推断 RIP 实际上限制了链孢霉的进化。有趣的是，一些真菌中发现了类似 RIP 的温和的基因组防卫系统。最显著的例子是减数分裂前诱导的甲基化（MIP），其过程是像 RIP 一样在受精和核融合之间检测相连和非相连的序列重复，但是它仅仅发生 DNA 甲基化，造成目标序列失活，但不发生突变（Rossignol and Faugeron 1994）。

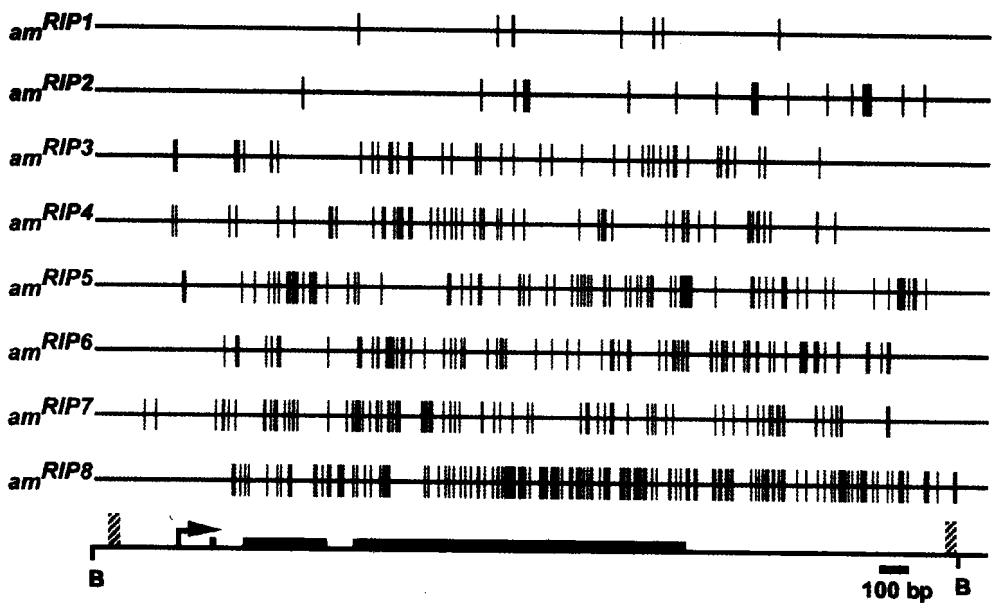



图 6-12 由 RIP 导致的突变以及八个 *am* 等位基因中的甲基化状态

垂直线代表突变。黑色等位基因不被甲基化。蓝色的等位基因最初是甲基化的，但是通过 5-azacytidine 诱导、克隆或者基因替换失去甲基化后，不能被从头甲基化。红色等位基因不但最初是甲基化的，而且也可以启动从头甲基化（引自 Singer et al. 1995）。

## 2.3 研究 RIP 的遗迹有助于了解 DNA 甲基化的调控

### 非经典的维持性甲基化

一个单独的 DNA 甲基转移酶 DIM-2 负责所有检测到的 DNA 甲基化，这和当初在链孢霉中发现非对称位点的甲基化一样令人吃惊，因为此前没有一个被鉴定的 DMT 能够甲基化多种上下文序列中的胞嘧啶。一个显而易见的重要问题是，是否非对称位点的甲基化能反映出某种特定序列有诱导从头甲基化的潜能？早期转化试验支持了这种可能性，被去掉甲基化的序列（如通过克隆）被转入营养细胞后重新获得了正常的甲基化水平。但当检测由 RIP 产生的八个 *am* 基因等位基因诱导从头甲基化的能力时，问题随之而来（Singer et al. 1995）。那些含有较少突变的 RIP 产物（图 6-12，$am^{RIP3}$ 和 $am^{RIP4}$）没有被从头甲基化，即便是在其正常所在的基因组，这说明我们观察到的甲基化是先建立的甲基化的传播。重要的是，正如在链孢霉中发现的其他甲基化一样，他们的甲基化并不限于对称的位点，也不显著的随时间蔓延，并且是“异质的”，即在一个克隆细胞群中甲基化残基的模式并不是均一的。因此，尽管依赖于先前在有性生殖阶段建立的甲基化（可能来自 RIP），但是这些甲基化并不由最初甲基化继承模型中预想的“维持性甲基化酶”来完成。值得注意的是 *Ascobolus* 中的 MIP 也会导致非均一的甲基化，这在真菌中提供了 DNA 甲基化传播的第一个证据（Rossignol and Faugeron 1994）。链孢霉进行维持性甲基化的能力被实验证实（Selker et al. 2002）。有趣的是，甲基化的传播

被发现是序列专一性的（即，它并不针对所有的序列），这为维持性甲基化概念添加了一个新的方面。虽然猜想了一系列能够导致 DNA 甲基化传播的潜在方案，但是在链孢霉中，实际的机制仍然未知。原则上，非对称位点甲基化的维持可以依赖于邻近的对称位点甲基化，但是观察到包括 CpG 位点的非均一性甲基化，使这种推测显得不太可能。涉及甲基化 DNA 结合蛋白的反馈机制可以产生预甲基化依赖性的甲基化，即，维持性甲基化。如下文讨论的，链孢霉和其他生物体中的发现表明组蛋白修饰调控了 DNA 甲基化，提出了组蛋白在甲基化维持中起作用的可能性。

### DNA 甲基化涉及组蛋白

组蛋白在 DNA 甲基化中的作用的第一个线索来自以下发现，用组蛋白去乙酰化酶抑制剂 TSA 处理链孢霉会降低某些染色质区域的甲基化水平（Selker 1998）。TSA 引起的去甲基化的选择性可能反映了对组蛋白乙酰转移酶的不同接近程度，但是这没有被完全证实（Selker et al. 2002）。*dim-5* 突变体中 DNA 甲基化，如 *dim-2* 中一样，被完全丢失，这无疑证明了染色质和 DNA 甲基化之间的关系。SET 结构域蛋白 DIM-5 被发现是专一性催化 H3K9 三甲基化的甲基转移酶（Tamaru and Selker 2001；Tamaru et al. 2003）。有两个证据可以证明组蛋白 H3 是 DIM-5 的底物：①H3 第九位赖氨酸被别的氨基酸取代引起 DNA 甲基化的丢失；②H3K9 三甲基化被发现专一性的存在于甲基化的染色质区域。

组蛋白甲基化可以调控 DNA 甲基化的发现（至少在链孢霉中）提出了两个重要的问题：①什么告诉 DIM-5 去甲基化哪个核小体？②什么阅读三甲基化标记然后把信息传给 DMT，DIM-2？调控 DIM-5 的因子还没有被确认，但有很强的证据表明其可被一个或多个识别 RIP 产物的蛋白调控，也被其催化位点周围氨基酸残基的修饰水平调控（Selker et al. 2002）。后者使 DIM-5 能够整合该特定区域 DNA 是否应该被甲基化的信息，这也为 TSA 阻止某些区域的 DNA 甲基化提供了一个可能的解释。

其他系统的研究结果导致发现了在链孢霉中是什么来阅读 H3K9 三甲基化标记。一个最早在果蝇中发现的蛋白 HP1（见第 5 章）可以在体外和甲基化的 H3K9 结合，这促使了在链孢霉中寻找 HP1 的同源基因。一个可能的同源基因已经被找到，并且通过基因敲除来证明它是否参与 DNA 甲基化（Freitag et al. 2004a）。事实上这个叫 *hpo*（HPone）的基因的确为 DNA 甲基化所必需。为了研究链孢霉中的 HP1 是否阅读 DIM-5 产生的标记，在野生型菌株和 *dim-5* 突变的菌株中检测它的亚细胞定位。在野生型中，HP1-GFP 定位在异染色质位点，但是这种定位在 *dim-5* 突变体中消失，这证明 HP1 是被 DIM-5 催化的三甲基化标记所招募的。

裂殖酵母中 RNA 干扰（RNAi）在异染色质形成和维持中起重要作用，这让人思考在链孢霉中 RNAi 机制是否参与 HP1 定位和 DNA 甲基化。一些参与 RNAi 的基因在链孢霉中有同源基因（Borkovich et al. 2004）。对三个 RNA 依赖的 RNA 聚合酶基因、Dicer 基因和其他推测的 RNAi 基因的突变体的研究并不能证明在链孢霉中 RNAi 参与 H3K9 甲基化、异染色质形成和 DNA 甲基化（Chicas et al. 2004；Freitag et al. 2004c）。但是，如下所述，在链孢霉中 RNAi 至少参与了另两种表观遗传学沉默机制，

平抑和减数分裂沉默。

## 2.4 平抑

在链孢霉转化技术建立起来后，一些实验室的研究人员注意到相当一部分（如30%）转化株中转入的基因会发生沉默，更令人吃惊的是，和转入基因序列同源的基因也会发生沉默。Macino 实验室命名无性生殖阶段的沉默为“平抑”（图 6-13），该实验

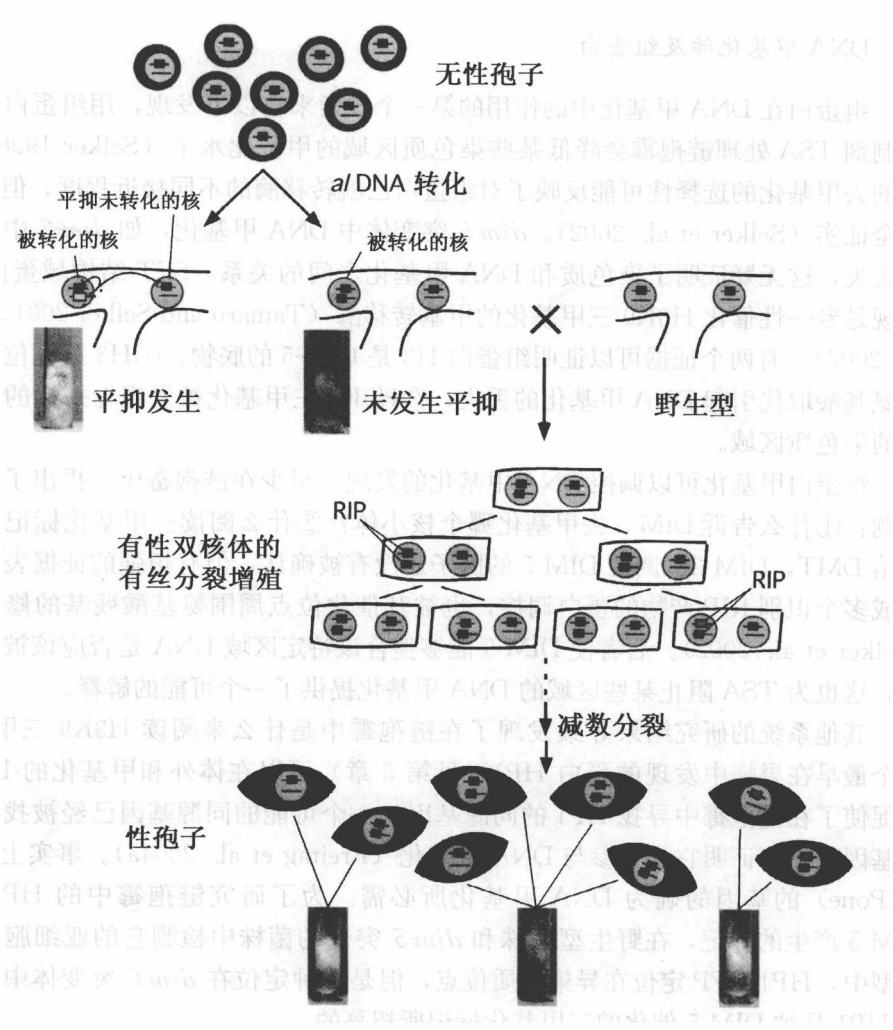



图 6-13 平抑

为简单起见，图中仅仅表示出七条染色体中的两条（灰色的圆代表核，其中直线代表染色体）。上方染色体中的深橙色长方形表示内源的 *albino* 基因（*al*），另一个（深橙色或黄色）矩形代表转入的异位的 *al* 序列。如图所示，既然转化的细胞一般是多核的，那么转化株也就一般是异核体。不论转化的 DNA 是否包含完整的编码区，在一些转化株中它可以通过一种未知的反式作用因子（从用黄色矩形表示的转化基因上发散出的红线）来沉默（“平抑”）被转化和未被转化的核中的天然 *al* 基因，结果是在某些转化株中出现浅色或 *albino* 缺陷（$Al^-$）的组织。

室到目前为止在这方面做出最多贡献（Pickford et al. 2002）。平抑被认为类似于植物中的共抑制和转录后基因沉默（PTGS），可视化的标记基因（如编码参与类胡萝卜素生物合成的白化基因 *albino*）的平抑最为清晰可见。有趣的是，基因对平抑的敏感性是不同的。对于那些敏感的基因，平抑更容易发生于那些带有多个拷贝并且紧密排列的转化 DNA 的转化株中。在链孢霉的菌丝中，核是自由流动的，异型核细胞中遗传上不同的核存在于同一细胞质中。很容易证明平抑是“显性的”；例如，转入的核可以使周围核中的同源序列沉默（Cogoni et al. 1996）。这暗示细胞质中存在沉默因子，或许是某种 RNA。有证据支持这种可能性，对参与平抑的基因的确认揭示出平抑和其他系统中发现的 RNAi 有很大关系（Pickford et al. 2002）。具体涉及的基因有 *qde-1*、*qde-2*、*qde-3*、*dcl-1* 和 *dcl-2*，分别编码一个 RNA 依赖的 RNA 聚合酶、一个被认为参与小干扰 RNA（siRNA）引导的 RNA 剪切的类“Argonaute”蛋白（Baumberger and Baulcombe 2005）、一个推测的类 RecQ 的 DNA 解旋酶（qde-3）和两个“Dicer”（被认为参与 siRNA的产生）（Catalanotto et al. 2004）。虽然缺乏关键的细节，但是正在完善中的模型是：在某种情况下，转入的 DNA 产生了非正常的转录物，以某种方式启动了 RNAi 机器，导致同源 mRNA 的降解。虽然转入的 DNA 经常发生甲基化，但是平抑的发生既不需要 DNA 甲基化酶 DIM-2 也不需要 H3K9 甲基转移酶 DIM-5（Cogoni et al. 1996；Chicas et al. 2005）。

## 2.5　非配对 DNA 导致的减数分裂沉默

MSUD 是由 Metzenberg 及其同事（Aramayo and Metzenberg 1996；shiu et al. 2001；shiu and Metzenberg 2002）在研究一个奇怪现象（子囊孢子成熟相关的 *Asm-1* 基因的缺失突变在功能上是显性的）时发现的，它是最晚被加入链孢霉沉默机制列表中的。图 6-14 所示的一系列完美的实验得出了以下结论：减数分裂前期非配对的序列会引起相同或近似相同的序列的沉默。*Asm-1* 的缺失是显性的，对此可能的解释是剩下的一个基因不能产生足够的基因产物，但是又发现一个异位的 *Asm-1* 基因并不能补救缺陷的表型（图 6-14c）。相反，如果序列以配对的方式存在的话，只要一个有功能的拷贝就能在减数分裂中发挥作用，例如，当它的另一个等位基因含有突变从而不能产生有功能的产物时（图 6-14d）。天然的两个等位基因都被缺失，但是存在配对的异位等位基因时，可以检测到蛋白的正常表达，这表明基因缺失本身并不造成破坏而异位的基因是可以发挥作用的（图 6-14e）。在可以进行 DNA 甲基化的菌株中，经过 RIP 作用的等位基因可以引起 MSUD，在 *dim-2* 突变的菌株中，则不能，这说明这些等位基因中高频率的 DNA 甲基化会阻止配对（比较图 6-14 中 d 和 f）（Pratt et al. 2004）。（偶然的，这个发现为 DIM-2 在链孢霉有性生殖阶段也发挥作用提供了第一个证据）RIP 造成的高密度突变的等位基因被发现和缺失一样是显性的（图 6-14g）（Shiu et al. 2001；Shiu and Metzenberg 2002；Pratt et al. 2004），这符合他们无法和野生型等位基因配对的猜测。为了弄清楚 MSUD 是由于不存在配对的等位基因还是由于存在不配对的等位基因，研究人员分析了一个杂交菌株，其减数分裂的核会含有一个基因的三个拷贝，两个野生型等位基因（配对的），一个异位的拷贝（非配对的）。沉默被观察到，这说明 MSUD

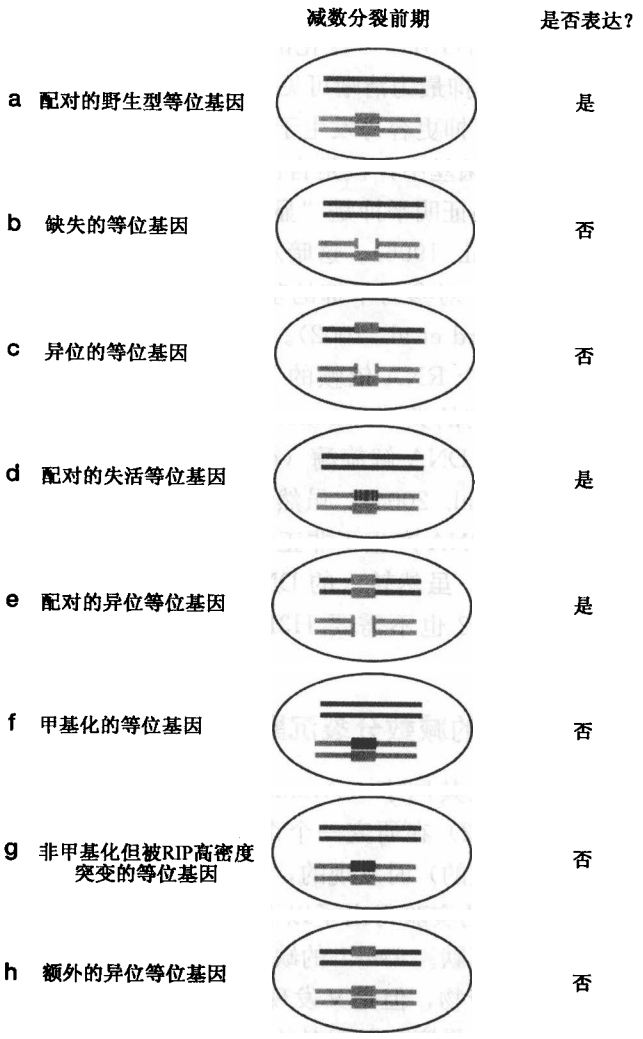



图 6-14 非配对 DNA 导致的减数分裂沉默

由 Metzenberg 和 Aramayo 实验室（Aramayo and Metzenberg 1996；Shiu et al. 2001）设计的检测方法，此现象也是在这些实验室被确认的。这里仅表示出七条染色体中的两条，矩形框表示一个一般在有丝分裂中起作用的基因。竖线表示突变，红色表示 DNA 甲基化。缺口表示基因缺失。

是由非配对的等位基因引起的，而不是由于不存在配对的等位基因（图 6-14h）。在含有编码带 GFP 标签的蛋白的基因的杂合子中，能在细胞水平上观察到 MSUD，如图 6-15 所示。

通过寻找 MSUD 缺陷的突变体，发现了很多 MSUD 机器的成员。在 *asm-1* 不配对的情况下，筛选能通过杂交的突变体，发现了 *sad-1*（子囊显性抑制因子），其编码一个 RdRP（Shiu et al. 2001）。这说明在链孢霉中，MSUD 和平抑相关，也和 RNAi 有关。有趣的是，据预测链孢霉的基因组编码三个 RdRP（一个在平抑中需要，一个在 MSUD 中需要，一个功能未知）和两个类 Dicer 蛋白。了解 MSUD 的具体机制令人着迷。

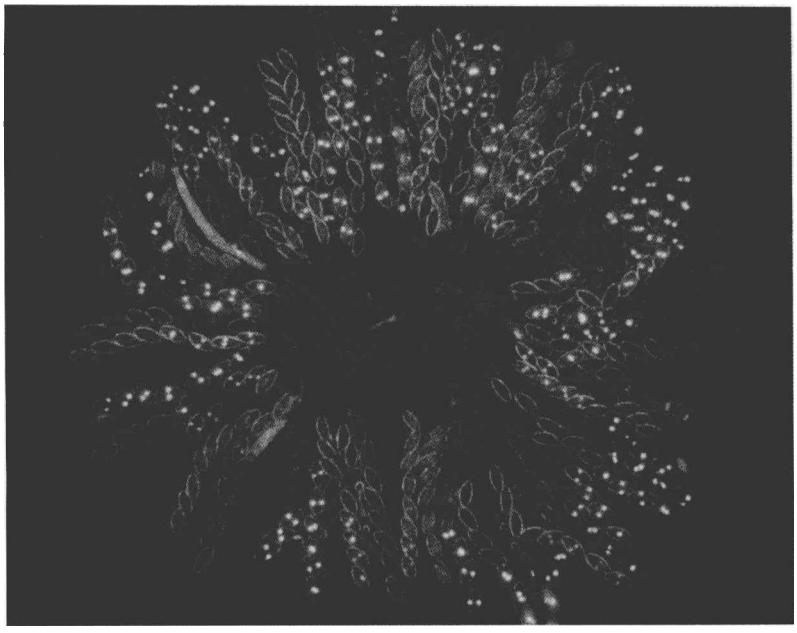


图 6-15　链孢霉

成熟中子囊（来自表达带 GFP 标签组蛋白 H1 转化株的杂合子代）的圆花窗样的荧光图像。每个子囊中有四个孢子有发光的核（Freitag et al. 2004b）（由斯坦福大学 N. B. Raju 惠赠）。

## 2.6　RIP、平抑和 MSUD 可能的功能和实际应用

RIP 似乎可用来限制“自私 DNA”的表达，如转座子，它们可以在基因组中产生自己的拷贝。和这种可能性相符的是，大量 RIP 的遗迹和其他物种中的转座子类似，大部分链孢霉的菌株缺少活跃的转座子（Galagan et al. 2003；selket et al. 2003）。然而，由于 RIP 只能在减数分裂前的双核细胞中发生，所以它既无法阻止营养细胞中新的（如水平得到的）转座子的蔓延，也不能阻止单拷贝转座子在减数分裂中的扩增。平抑和 MSUD 可以处理此类事件。虽然，如转入的类 LINE 转座子的增殖所证明的，平抑并不能完全抑制营养细胞中转座子的蔓延，但是它确实能部分的沉默此类转座子（Nolan et al. 2005）。有关 MSUD 的信息显示它能沉默减数分裂细胞中任何移位的序列，即便其在基因组中仅仅只存在一个拷贝（Shiu et al. 2001）。除了处理减数分裂中流浪的转座子，MSUD 似乎也在物种形成中扮演了重要角色，MSUD 缺陷的突变体解除了含有大量染色体扩增片段的菌株的不育性，并且使链孢霉可以和关系相近的种交配（Shiu et al. 2001）。

尽管 RIP、平抑和 MSUD 会为某些遗传学实验带来麻烦，但它们都已经被开发并用于研究。RIP 为在链孢霉中敲除基因提供了最简单的方法，也是产生“部分功能”突变体的常用方法。平抑被用于降低（不是完全缺失）基因表达，如 RNAi 在其他生物体中的应用一样。另外，MSUD 提供了一个简单的检测方法，可以用来检测某个基因是

否被减数分裂所需要；如果一个基因发生扩增的时候可以引起不育，而在能够阻止MSUD的突变体中这种不育表型能够恢复，那么基本可以预测这个基因在减数分裂中起重要作用。

除了假设在进化中起作用，以及在实验室被利用之外，RIP、平抑和MSUD或许在其他方面也发挥作用。例如，完整的生育力需要 *Sad-1*，所以可以认为减数分裂也直接或者间接的需要MSUD（Shiu and Metzenberg 2002）。RIP是非必需的，但是RIP产物在链孢霉基因组中的分布证明这些被废弃的转座子可以作为动粒装配的基础，就像裂殖酵母和其他生物体中的重复序列一样。染色体7的DNA序列分析表明链孢霉着丝粒周围的序列主要含有被RIP严重突变的转座子的遗迹（图6-16）。在链孢霉端粒周围也发现了RIP遗迹。有趣的是，在其他生物体异染色质中也发现了转座子和转座子遗迹，如果蝇（第5章）、哺乳动物（第17章）、植物（第9章）和其他的真菌。

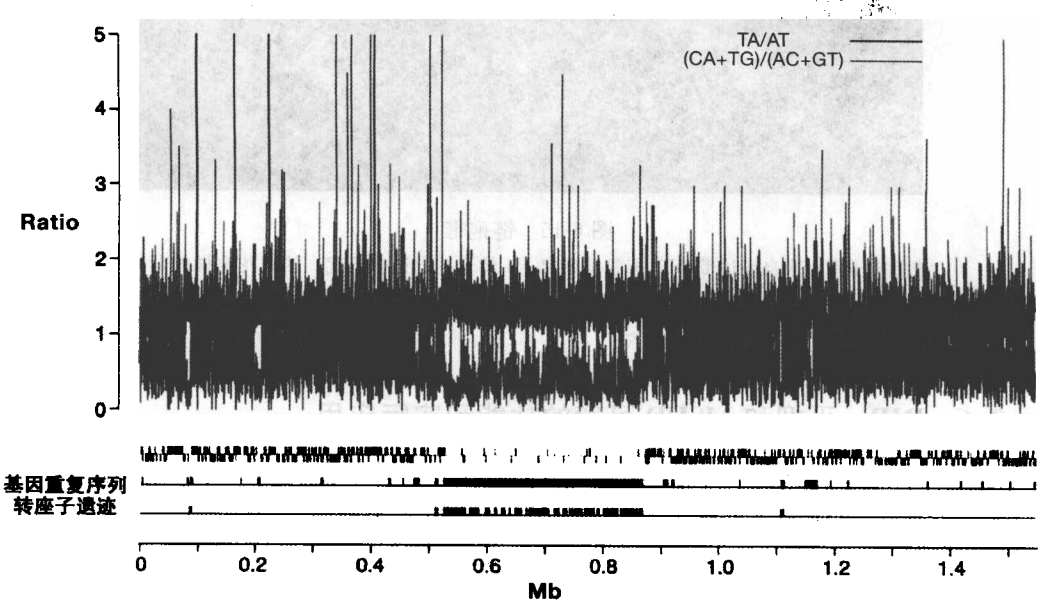



图6-16 链孢霉7号染色体区的组织结构

来自genomic sequence release 7.0(http://www.broad.mit.edu/annotation/fungi/neurospora_crass_7/index.html)的克隆叠连群249、255、21，除了第10叠连群的前400kb序列，所有序列均已被拼接在一起，拼接序列以200bp为单位进行“RIP指数”分析（TpA/ApT［蓝色］and CpA+TpG/ApC+GpT［红色］）（Galagan et al. 2003；Selker et al. 2003；Galagan and Selker 2004）。在包围着丝粒的标记（用遗传学方法被定位）之间，发现一个约360kb的区域含有高密度的被RIP失活的转座子元件（逆转座子遗迹用蓝色表示，DNA转座子用紫色表示）。这个约1.5Mb的连接片段中有383个被注释的基因（线上和线下表示基因的相反的定向），其中仅20个预测的短基因存在于预测的着丝粒区域。这些叠连群间的序列空隙大小未知（在图中0.5466、0.6956和0.9085Mb的位置）。

## 3. 总　结

裂殖酵母和链孢霉已经成为发现和阐明表观遗传学现象的强有力的系统。表观遗传

学领域仍处在发展初期，表观遗传学机制正逐步的为人所知。因此，不必对我们目前对表观遗传过程理解的深度和广度表示惊讶，如本章所描述的，不同生物体中有不同的过程。我们还不知道所描述的表观遗传机制有何种程度的普遍性。例如，裂殖酵母等生物体依赖 RNAi 来沉默基因和形成异染色质，然而其他生物如链孢霉却依赖 DNA 甲基化。现在已经明白，这两种模式真核生物既有重要的不同也有显著的相似。他们都利用组蛋白 H3K9 甲基化和 RNAi，而这些在酿酒酵母中都不存在。在这三种真菌中，只有链孢霉有 DNA 甲基化。另外需要指出的是，同一过程在不同生物体中可以有不同的功能。例如，在链孢霉中，RNAi 在平抑和减数分裂沉默中发挥作用，但不参与异染色质形成，然而在裂殖酵母中，RNAi 参与异染色质形成，还没发现其他的作用。最后需要注意的是，虽然有共同点，如在裂殖酵母和链孢霉中异染色质都和着丝粒有关，但也可能存在重要的不同。未来的一个重要目标是发现从一种生物体中得到的信息在其他生物体上的适用程度。对包括裂殖酵母和链孢霉在内的不同生物体表观遗传过程的进一步研究，会提供这些信息。我们预期在很多年内丰富多样的真菌会继续作为表观遗传研究的有用的系统。

## 致　谢

Robin C. Allshire 感谢 Alison L. Pidoux 和 Sharon A. White 分别在图 6-4 和图 6-5 中的帮助；感谢 Alison L. Pidoux 和 Elizabeth H. Bayne 对原稿的意见。Allshire 实验室的研究受到 Wellcome Trust、MRC 和 EC FP6 表观基因组学先进研究网络的资助。Robin C. Allshire 是 Wellcome Trust 研究员。Eric U. Selker 感谢 N. B. Raju 帮助收集链孢霉的图片，感谢 Robert L. Metazenberg 对原稿的意见。Selker 实验室的研究受到 NIH 美国公共健康服务项目 GM35690 和国家科学基金项目 MCB0131383 的资助。

马　奔　译
杨　鹏　校
朱　冰　审校

## 参 考 文 献

Allshire R.C., Javerzat J.P., Redhead N.J., and Cranston G. 1994. Position effect variegation at fission yeast centromeres. *Cell* **76:** 157–169.

Allshire R.C., Nimmo E.R., Ekwall K., Javerzat J.P., and Cranston G. 1995. Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev.* **9:** 218–233.

Aramayo R. and Metzenberg R.L. 1996. Meiotic transvection in fungi. *Cell* **86:** 103–113.

Ayoub N., Goldshmidt I., Lyakhovetsky R., and Cohen A. 2000. A fission yeast repression element cooperates with centromere-like sequences and defines a mat silent domain boundary. *Genetics* **156:** 983–994.

Bailis J.M., Bernard P., Antonelli R., Allshire R.C., and Forsburg S.L. 2003. Hsk1-Dfp1 is required for heterochromatin-mediated cohesion at centromeres. *Nat. Cell Biol.* **5:** 1111–1116.

Bannister A.J., Zegerman P., Partridge J.F., Miska E.A., Thomas J.O., Allshire R.C., and Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120–124.

Baum M., Ngan V.K., and Clarke L. 1994. The centromeric K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. *Mol. Biol. Cell* **5:** 747–761.

Baumberger N. and Baulcombe D.C. 2005. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc. Natl. Acad. Sci.* **102:** 11928–11933.

Bernard P., Maure J.F., Partridge J.F., Genier S., Javerzat J.P., and Allshire R.C. 2001. Requirement of heterochromatin for cohesion at centromeres. *Science* **294:** 2539–2542.

Bhat A. and Kasbekar D.P. 2001. Escape from repeat-induced point mutation of a gene-sized duplication in *Neurospora crassa* crosses that are heterozygous for a larger chromosome segment duplication. *Genetics* **157:** 1581–1590.

Borkovich K.A., Alex L.A., Yarden O., Freitag M., Turner G.E., Read

N.D., Seiler S., Bell-Pedersen D., Paietta J., Plesofsky N., et al. 2004. Lessons from the genome sequence of *Neurospora crassa*: Tracing the path from genomic blueprint to multicellular organism. *Microbiol. Mol. Biol. Rev.* **68:** 1–108.

Cam H.P., Sugiyama T., Chen E.S., Chen X., FitzGerald P.C., and Grewal S.I. 2005. Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* **37:** 809–819.

Catalanotto C., Pallotta M., ReFalo P., Sachs M.S., Vayssie L., Macino G., and Cogoni C. 2004. Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in *Neurospora crassa*. *Mol. Cell. Biol.* **24:** 2536–2545.

Chen E.S., Saitoh S., Yanagida M., and Takahashi K. 2003. A cell cycle-regulated GATA factor promotes centromeric localization of CENP-A in fission yeast. *Mol. Cell* **11:** 175–187.

Chicas A., Cogoni C., and Macino G. 2004. RNAi-dependent and RNAi-independent mechanisms contribute to the silencing of RIPed sequences in *Neurospora crassa*. *Nucleic Acids Res.* **32:** 4237–4243.

Chicas A., Forrest E.C., Sepich S., Cogoni C., and Macino G. 2005. Small interfering RNAs that trigger posttranscriptional gene silencing are not required for the histone H3 Lys9 methylation necessary for transgenic tandem repeat stabilization in *Neurospora crassa*. *Mol. Cell. Biol.* **25:** 3793–3801.

Clarke, L. and M.P. Baum. 1990. Functional analysis of a centromere from fission yeast: A role for centromere-specific repeated DNA sequences. *Mol. Cell. Biol.* **10:** 1863–1872.

Cleveland D.W., Mao Y., and Sullivan K.F. 2003. Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. *Cell* **112:** 407–421.

Cogoni C., Irelan J.T., Schumacher M., Schmidhauser T.J., Selker E.U., and Macino G. 1996. Transgene silencing of the *al-1* gene in vegetative cells of *Neurospora* is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. *EMBO J.* **15:** 3153–3163.

Cowieson N.P., Partridge J.F., Allshire R.C., and McLaughlin P.J. 2000. Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr. Biol.* **10:** 517–525.

Davis R.H. 2000. Neurospora: *Contributions of a model organism*. Oxford University Press, New York.

Djupedal I., Portoso M., Spahr H., Bonilla C., Gustafsson C.M., Allshire R.C., and Ekwall K. 2005. RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev.* **19:** 2301–2306.

Egel R. 2004. *The molecular biology of* Schizosaccharomyces pombe: *Genetics, genomics and beyond*. Springer-Verlag, Heidelberg, Germany.

Ekwall K., Javerzat J.P., Lorentz A., Schmidt H., Cranston G., and Allshire R. 1995. The chromodomain protein Swi6: A key component at fission yeast centromeres. *Science* **269:** 1429–1431.

Ekwall K., Nimmo E.R., Javerzat J.P., Borgstrom B., Egel R., Cranston G., and Allshire R. 1996. Mutations in the fission yeast silencing factors *clr4*$^+$ and *rik1*$^+$ disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. *J. Cell Sci.* **109:** 2637–2648.

Ekwall K., Olsson T., Turner B.M., Cranston G., and Allshire R.C. 1997. Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* **91:** 1021–1032.

Freitag M., Hickey P.C., Khlafallah T.K., Read N.D., and Selker E.U. 2004a. HP1 is essential for DNA methylation in *Neurospora*. *Mol. Cell* **13:** 427–434.

Freitag M., Hickey P.C., Raju N.B., Selker E.U., and Read N.D. 2004b. GFP as a tool to analyze the organization, dynamics and function of nuclei and microtubules in *Neurospora crassa*. *Fungal Genet. Biol.* **41:** 897–910.

Freitag M., Lee D.W., Kothe G.O., Pratt R.J., Aramayo R., and Selker E.U. 2004c. DNA methylation is independent of RNA interference in *Neurospora*. *Science* **304:** 1939.

Freitag M., Williams R.L., Kothe G.O., and Selker E.U. 2002. A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. *Proc. Natl. Acad. Sci.* **99:** 8802–8807.

Fukagawa T., Nogami M., Yoshikawa M., Ikeno M., Okazaki T., Takami Y., Nakayama T., and Oshimura M. 2004. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat. Cell Biol.* **6:** 784–791.

Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D., Jaffe D., FitzHugh W., Ma L.J., Smirnov S., Purcell S., et al. 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422:** 859–868.

Galagan J.E. and Selker E.U. 2004. RIP: The evolutionary cost of genome defense. *Trends Genet.* **20:** 417–423.

Goshima G., Saitoh S., and Yanagida M. 1999. Proper metaphase spindle length is determined by centromere proteins Mis12 and Mis6 required for faithful chromosome segregation. *Genes Dev.* **13:** 1664–1677.

Grace Goll M. and Bestor T.H. 2005. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74:** 481–514.

Grewal S.I. and Klar A.J. 1997. A recombinationally repressed region between mat2 and mat3 loci shares homology to centromeric repeats and regulates directionality of mating-type switching in fission yeast. *Genetics* **146:** 1221–1238.

Hall I.M., Noma K., and Grewal S.I. 2003. RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast. *Proc. Natl. Acad. Sci.* **100:** 193–198.

Hall I.M., Shankaranarayana G.D., Noma K., Ayoub N., Cohen A., and Grewal S.I. 2002. Establishment and maintenance of a heterochromatin domain. *Science* **297:** 2232–2237.

Hayashi T., Fujita Y., Iwasaki O., Adachi Y., Takahashi K., and Yanagida M. 2004. Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* **118:** 715–729.

Hong E.-J.E., Villen J., Gerace E.L., Gygi S., and Moazed D. 2005. A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. *RNA Biol.* **2:** 106–111.

Irelan J.T., Gutkin G.I., and Clarke L. 2001. Functional redundancies, distinct localizations and interactions among three fission yeast homologs of centromere protein-B. *Genetics* **157:** 1191–1203.

Jia S., Noma K., and Grewal S.I. 2004. RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science* **304:** 1971–1976.

Jin Q.W., Pidoux A.L., Decker C., Allshire R.C., and Fleig U. 2002. The mal2p protein is an essential component of the fission yeast centromere. *Mol. Cell Biol.* **22:** 7168–7183.

Kanoh J., Sadaie M., Urano T., and Ishikawa F. 2005. Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Curr. Biol.* **15:** 1808–1819.

Kato H., Goto D.B., Martienssen R.A., Urano T., Furukawa K., and Murakami Y. 2005. RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* **309:** 467–469.

Kim H.S., Choi E.S., Shin J.A., Jang Y.K., and Park S.D. 2004. Regulation of Swi6/HP1-dependent heterochromatin assembly by cooperation of components of the mitogen-activated protein kinase pathway and a histone deacetylase Clr6. *J. Biol. Chem.* **279:** 42850–42859.

Kouzminova E.A. and Selker E.U. 2001. *Dim-2* encodes a DNA-methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J.* **20:** 4309–4323.

Kuhn R.M., Clarke L., and Carbon J. 1991. Clustered tRNA genes in *Schizosaccharomyces pombe* centromeric DNA sequence repeats. *Proc. Natl. Acad. Sci.* **88:** 1306–1310.

Measday V. and Hieter P. 2004. Kinetochore sub-structure comes to MIND. *Nat. Cell Biol.* **6:** 94–95.

Mellone B.G., Ball L., Suka N., Grunstein M.R., Partridge J.F., and Allshire R.C. 2003. Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. *Curr Biol* **13:** 1748–1757.

Motamedi M.R., Verdel A., Colmenares S.U., Gerber S.A., Gygi S.P., and Moazed D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119:** 789–802.

Nakagawa H., Lee J.K., Hurwitz J., Allshire R.C., Nakayama J., Grewal S.I., Tanaka K., and Murakami Y. 2002. Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications. *Genes Dev.* **16:** 1766–1778.

Nakayama J., Klar A.J., and Grewal S.I. 2000. A chromodomain protein, Swi6, performs imprinting functions in fission yeast during mitosis and meiosis. *Cell* **101:** 307–317.

Nakayama J., Rice J.C., Strahl B.D., Allis C.D., and Grewal S.I. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292:** 110–113.

Ngan V.K. and Clarke L. 1997. The centromere enhancer mediates centromere activation in *Schizosaccharomyces pombe*. *Mol. Cell. Biol.* **17:** 3305–3314.

Nimmo E.R., Cranston G., and Allshire R.C. 1994. Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *EMBO J.* **13:** 3801–3811.

Nimmo E.R., Pidoux A.L., Perry P.E., and Allshire R.C. 1998. Defective meiosis in telomere-silencing mutants of *Schizosaccharomyces pombe*. *Nature* **392:** 825–828.

Nolan T., Braccini L., Azzalin G., De Toni A., Macino G., and Cogoni C. 2005. The post-transcriptional gene silencing machinery functions independently of DNA methylation to repress a LINE1-like retrotransposon in *Neurospora crassa*. *Nucleic Acids Res.* **33:** 1564–1573.

Noma K., Allis C.D., and Grewal S.I. 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* **293:** 1150–1155.

Noma K., Sugiyama T., Cam H., Verdel A., Zofall M., Jia S., Moazed D., and Grewal S.I. 2004. RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* **36:** 1174–1180.

Nonaka N., Kitajima T., Yokobayashi S., Xiao G., Yamamoto M., Grewal S.I., and Watanabe Y. 2002. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat. Cell Biol.* **4:** 89–93.

Partridge J.F., Borgstrom B., and Allshire R.C. 2000. Distinct protein interaction domains and protein spreading in a complex centromere. *Genes Dev.* **14:** 783–791.

Partridge J.F., Scott K.S., Bannister A.J., Kouzarides T., and Allshire R.C. 2002. cis-acting DNA from fission yeast centromeres mediates histone H3 methylation and recruitment of silencing factors and cohesin to an ectopic site. *Curr. Biol.* **12:** 1652–1660.

Perkins D.D., Margolin B.S., Selker E.U., and Haedo S.D. 1997. Occurrence of repeat induced point mutation in long segmental duplications of *Neurospora*. *Genetics* **147:** 125–136.

Perkins D.D., Metzenberg R.L., Raju N.B., Selker E.U., and Barry E.G. 1986. Reversal of a *Neurospora* translocation by crossing over involving displaced rDNA, and methylation of the rDNA segments that result from recombination. *Genetics* **114:** 791–817.

Perkins D.D., Radford A., and Sachs M.S. 2001. *The* Neurospora *compendium; Chromosomal loci*. Academic Press, San Diego, California.

Pickford A.S., Catalanotto C., Cogoni C., and Macino G. 2002. Quelling in *Neurospora crassa*. *Adv. Genet.* **46:** 277–303.

Pidoux A.L. and Allshire R.C. 2004. Kinetochore and heterochromatin domains of the fission yeast centromere. *Chromosome Res.* **12:** 521–534.

———. 2005. The role of heterochromatin in centromere function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360:** 569–579.

Pidoux A.L., Richardson W., and Allshire R.C. 2003. Sim4: A novel fission yeast kinetochore protein required for centromeric silencing and chromosome segregation. *J. Cell Biol.* **161:** 295–307.

Polizzi C. and Clarke L. 1991. The chromatin structure of centromeres from fission yeast: Differentiation of the central core that correlates with function. *J. Cell Biol.* **112:** 191–201.

Pratt R.J., Lee D.W., and Aramayo R. 2004. DNA methylation affects meiotic trans-sensing, not meiotic silencing, in *Neurospora*. *Genetics* **168:** 1925–1935.

Raju N.B. 1980. Meiosis and ascospore genesis in *Neurospora*. *Eur. J. Cell Biol.* **23:** 208–223.

Reinhart B.J. and Bartel D.P. 2002. Small RNAs correspond to centromere heterochromatic repeats. *Science* **297:** 1831.

Rossignol J.-L. and Faugeron G. 1994. Gene inactivation triggered by recognition between DNA repeats. *Experientia* **50:** 307–317.

Sadaie M., Iida T., Urano T., and Nakayama J. 2004. A chromodomain protein, Chp1, is required for the establishment of heterochromatin in fission yeast. *EMBO J.* **23:** 3825–3835.

Scott K.C., Merrett S.L., and Willard H.F. 2006. A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains. *Curr. Biol.* **16:** 119–129.

Selker E.U. 1990. Premeiotic instability of repeated sequences in *Neurospora crassa*. *Annu. Rev. Genet.* **24:** 579–613.

———. 1998. Trichostatin A causes selective loss of DNA methylation in *Neurospora*. *Proc. Natl. Acad. Sci.* **95:** 9430–9435.

Selker E.U., Cambareri E.B., Jensen B.C., and Haack K.R. 1987. Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell* **51:** 741–752.

Selker E.U., Freitag M., Kothe G.O., Margolin B.S., Rountree M.R., Allis C.D., and Tamaru H. 2002. Induction and maintenance of nonsymmetrical DNA methylation in *Neurospora*. *Proc. Natl. Acad. Sci.* (suppl. 4) **99:** 16485–16490.

Selker E.U., Tountas N.A., Cross S.H., Margolin B.S., Murphy J.G., Bird A.P., and Freitag M. 2003. The methylated component of the *Neurospora crassa* genome. *Nature* **422:** 893–897.

Shankaranarayana G.D., Motamedi M.R., Moazed D., and Grewal S.I. 2003. Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. *Curr. Biol.* **13:** 1240–1246.

Shiu P.K. and Metzenberg R.L. 2002. Meiotic silencing by unpaired DNA: Properties, regulation and suppression. *Genetics* **161:** 1483–1495.

Shiu P.K., Raju N.B., Zickler D., and Metzenberg R.L. 2001. Meiotic silencing by unpaired DNA. *Cell* **107:** 905–916.

Sigova A., Rhind N., and Zamore P.D. 2004. A single Argonaute protein mediates both transcriptional and posttranscriptional silencing in *Schizosaccharomyces pombe*. *Genes Dev.* **18:** 2359–2367.

Singer M.J., Marcotte B.A., and Selker E.U. 1995. DNA methylation associated with repeat-induced point mutation in *Neurospora crassa*. *Mol. Cell. Biol.* **15:** 5586–5597.

Steiner N.C. and Clarke L. 1994. A novel epigenetic effect can alter centromere function in fission yeast. *Cell* **79:** 865–874.

Steiner N.C., Hahnenberger K.M., and Clarke L. 1993. Centromeres of the fission yeast *Schizosaccharomyces pombe* are highly variable genetic loci. *Mol. Cell. Biol.* **13:** 4578–4587.

Sugiyama T., Cam H., Verdel A., Moazed D., and Grewal S.I. 2005. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci.* **102:** 152–157.

Takahashi K., Chen E.S., and Yanagida M. 2000. Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast. *Science* **288:** 2215–2219.

Takahashi K., Murakami S., Chikashige Y., Funabiki H., Niwa O., and Yanagida M. 1992. A low copy number central sequence with strict symmetry and unusual chromatin structure in fission yeast centromere. *Mol. Biol. Cell* **3:** 819–835.

Takahashi K., Murakami S., Chikashige Y., Niwa O., and Yanagida M. 1991. A large number of tRNA genes are symmetrically located in fission yeast centromeres. *J. Mol. Biol.* **218:** 13–17.

Takahashi K., Takayama Y., Masuda F., Kobayashi Y., and Saitoh S. 2005. Two distinct pathways responsible for the loading of CENP-A to centromeres in the fission yeast cell cycle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360:** 595–607.

Tamaru H. and Selker E.U. 2001. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414:** 277–283.

———. 2003. Synthesis of signals for de novo DNA methylation in *Neurospora crassa*. *Mol.Cell. Biol.* **23:** 2379–2394.

Tamaru H., Zhang X., McMillen D., Singh P.B., Nakayama J., Grewal S.I., Allis C.D., Cheng X., and Selker E.U 2003. Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. *Nat. Genet.* **34:** 75–79.

Thon G. and Verhein-Hansen J. 2000. Four chromo-domain proteins of

*Schizosaccharomyces pombe* differentially repress transcription at various chromosomal locations. *Genetics* **155:** 551–568.

Volpe T., Schramke V., Hamilton G.L., White S.A., Teng G., Martienssen R.A., and Allshire R.C. 2003. RNA interference is required for normal centromere function in fission yeast. *Chromosome Res.* **11:** 137–146.

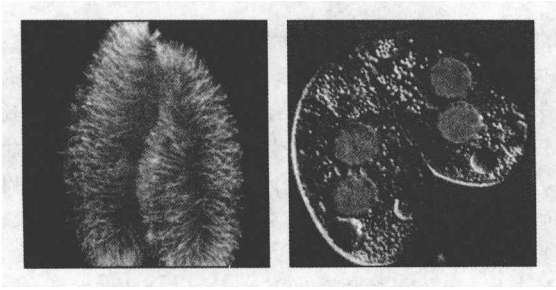
Volpe T.A., Kidner C., Hall I.M., Teng G., Grewal S.I., and Martienssen R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297:** 1833–1837.

Watters M.K., Randall T.A., Margolin B.S., Selker E.U., and Stadler D.R. 1999. Action of repeat-induced point mutation on both strands of a duplex and on tandem duplications of various sizes in *Neurospora*. *Genetics* **153:** 705–714.

Wilkinson C.R., Bartlett R., Nurse P., and Bird A.P. 1995. The fission yeast gene pmt1+ encodes a DNA methyltransferase homologue. *Nucleic Acids Res.* **23:** 203–210.

Yamada T., Fischle W., Sugiyama T., Allis C.D., and Grewal S.I. 2005. The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Mol. Cell* **20:** 173–185.

# 第 7 章



# 纤毛动物的表观遗传学

Eric Meyer[1] and Douglas L. Chalker[2]

[1]*Laboratoire de Genetique Moleculaire, CNRS UMR 8541 Ecole Normale Superieure, 75005 Paris, France*
[2]*Deptartment of Biology, Washington University, St. Louis, Missouri 63130*

## 本章目录

1. 纤毛动物：含两套不同基因组的单细胞生物
2. 接合：相互受精揭示非孟德尔遗传
3. 纤毛动物的胞质遗传
4. 皮质模式：结构继承的一个例子
5. 大核与小核：活性染色质对沉默染色质的模型
   - 5.1 大核与小核组蛋白的分离表明了不同组蛋白的不同功能
   - 5.2 染色质修饰与染色质的活性状态相关
6. 整个基因组范围的重排发生于大核发育时期
   - 6.1 内源 DNA 缺失：精确和非精确机制
   - 6.2 染色体片段化
7. 基因组重排的机制
8. 纤毛动物中同源依赖性的基因沉默
   - 8.1 转基因诱导的基因沉默
   - 8.2 基因沉默是由双链 RNA 诱导的
9. 基因组重排是同源依赖机制指导的
   - 9.1 实验诱导大核发育中的特异性缺失
10. 重排模式是由生殖细胞和体细胞基因组的比较而决定的
    - 10.1 d48 为例：其他重排方式的表观遗传
    - 10.2 实验诱导 DNA 缺失产生的表观遗传
    - 10.3 引入小核基因组的外源序列的“自发性”移除
    - 10.4 可继承的大核 DNA 缺失的实验补救
    - 10.5 实验性抑制发育中大核的 IES 移除
11. 由 RNA 干扰介导的整个基因组的跨核比较
    - 11.1 DNA 移除与短 RNA 的关系
    - 11.2 在草履虫中将 RNA 从母体转移到合子大核中
12. 结论：DNA 移除是基因组防御机制
13. 未来纤毛动物研究对表观遗传学的贡献

参考文献

## 概要

任何人在显微镜下观察纤毛动物都会被这些复杂的小生命所吸引，它们使用类似毛发的纤毛来游泳、觅食和寻找配偶。它们之中的无性生殖细胞通过简单的二分裂方式繁殖，但是它们也会周期性地与配偶交配，或者在某些种类中，能通过自交的方式繁殖，这就导致有性后代出现不同的基因型。这类单细胞的真核生物与众不同的地方在于，它们保持拥有两套功能不同的基因组，被同一胞质中的不同的核携带。两个中较小的一个称为小核，包含生殖相关的基因组。在生长的过程，小核上的基因是转录沉默的，但是却保存了传递给有性后代的遗传信息。较大的大核行使体细胞功能，它负责几乎所有的基因表达，因而对细胞的表型具有决定性作用。在每次无性分裂周期的末期，大核就被丢弃，新的大核由生殖细胞中分化出来。在大核的发育过程中，大量的 DNA 重排产生了全新的、待表达的基因组。生殖细胞基因组的一部分，包括所有的被视为“垃圾”的 DNA 重复序列都被移除，剩下的基因都是生物整个生命周期所必需的，这些基因被大量扩增，以多倍形式存在于基因组中。

纤毛动物这种遗传的特异性使它们成为发现和理解表观遗传机制非常有用的模式生物。在某些物种中，由同一细胞有性生殖分裂而成的两个后代开始具有相同的基因组，但是它们的体细胞核在亲本细胞的基因组不同时会出现区别。这使得我们可以对遗传性状进行很方便的检测，而这些性状本身并不是由核基因组所决定的。对草履虫进行的遗传实验，尤其是 Tracy Bonneborn 的实验，为真核生物中非孟德尔遗传现象的存在提供了最早的描述。在那些遗传组成上相同的同胞细胞却在某些特殊性状上呈现出不同的例子中，一部分是细胞质遗传的真正例子（细胞器 DNA 的母系遗传），而另一些却是属于不同的机理，例如某个体接合型的继承。后代的接合型不是由它自己的基因型决定的，而是由其亲本细胞先存在的接合型决定，后代的体细胞核正是在亲本细胞中发育的。简单地说，不同的环境（亲本的细胞质）会指导不同形状的表达，即使拥有相同的 DNA 组成也不例外。这就是表观遗传的标记。

纤毛生物与众不同的遗传结构也暗示了它不同核中同源的序列的不同调控机理。早期的研究关注于解析为何生殖细胞基因保持转录沉默，而体细胞基因组可以活跃转录。基因表达状态的区分为研究者提供了一个机会，即研究染色质蛋白和它们在表观遗传调控中的修饰作用。它们可以协调特定的组蛋白及其转录活性的修饰与细胞周期阶段的关系。例如，比较嗜热四膜虫生殖细胞和体细胞核的染色质蛋白，第一次发现了组蛋白的不同。另外，新的染色质调控子，例如第一个组蛋白乙酰转移酶（histone acetyltransferase，HAT）也是在这种纤毛动物中发现的，该发现归功于纤毛动物的一个优势：只有大核具有乙酰化的组蛋白。

尽管纤毛动物的遗传结构看上去并不典型，但是其隐藏的很多机制却在很多真核生物的表观遗传调控中广泛使用，例如 RNA 干扰在整个基因组重排中的作用。这些重排的程度和形式在不同的纤毛动物物种间具有很大的差异，但是一个共同的特点是它们通常都会移除类似转座子的元件以及其他的 DNA 重复序列。在草履虫和四膜虫中，短 RNA 是在生殖细胞减数分裂时由其基因组产生的。这些小 RNA 的发现，连同确定 Argonaute 及 Dicer 同源物是 DNA 重排所必需的事实暗示可能包含一个类似 RNA 干扰的机制。小 RNA 被认为作用于同源序列的组蛋白 H3 第 9 位赖氨酸，使得它们带上准备被移除的标记。因此，纤毛动物 DNA 的重排从机理上说与广泛使用的 RNA 介导的异染色质的形成是相似的。利用的 RNA 干扰的机制移除转座元件更进一步强调了该通路的重要性，因为很可能是基因组防御系统的组成部分。另外，很多实验表明 DNA 的重排模式并不是由生殖细胞基因组严格决定的，而是受调控的，至少一部分与亲本细胞体细胞基因组原有的重排相关。这表明在核分化的过程中生殖细胞和体细胞的基因组有比较，而这种比较可能就是生殖细胞和体细胞 RNA 依赖同源性的相互作用介导的。对这个过程更完整的理解毫无疑问会为 RNA 在基因组中表观遗传作用提供新的视角。

## 1. 纤毛动物：含两套不同基因组的单细胞生物

纤毛动物出现于一亿年前，包含一个单演化系，是最早用于遗传学研究的单细胞真核生物之一（Philippe et al. 2000）。在 20 世纪 30 年代末期，在 T. M Sonneborn 发现双小核草履虫的配偶型前（Sonneborn 1937），T. H. Morgan 提出的遗传的染色体理论就受到很多科学家的质疑，尤其是胚胎学家（历史详情参见第 2 章）。胚胎学家无法想象像基因这样的静态物质是如何成为遗传的唯一基础，他们相信细胞质一定包含于遗传的过程中，只要协调基因的行为即可（Harwood 1985）。然而，主流的遗传学家大量关注于基因的行为，Sonneborn 的早期遗传分析表明很多可继承的特征的传递是无法用孟德尔定律完全解释的。由于纤毛生物研究的特殊性，在该领域研究中发现了第一个细胞质参与遗传的例子，并且仍将对表观遗传的机制提供新的视角。

纤毛生物最与众不同的特征之一是其核二型性：每个细胞包含两种核，功能和结构都不相同。二倍体的小核在无性繁殖中是保持转录沉默的，但是却包含了生殖细胞的基因组。这些核会进行减数分裂，产生配子核并依照孟德尔定律将基因组传递给有性后代（图 7-1）。相反，多倍体的大核在无性繁殖期负责基因的表达，所以它们决定细胞的表型，但是在性发育过程中丢失，这与人类的体细胞性质类似（图 7-1）。不同物种中每种核的数目差异很大。例如，双小核草履虫有两个小核和一个大核，而嗜热四膜虫只有一个大核和一个小核。

大核与小核的分裂有不同的机制。小核通过传统的封闭的减数分裂完成分裂。而大核则是通过一种机制尚未明确的无丝分裂的方式完成分裂，其整个过程不包含纺锤丝的形成，也无法观察到非中心粒区域染色质的浓缩。合成 DNA 以后，大核很简单地分为大致相同的两半。现在并没有发现任何机制表明大核的染色体是均匀分配到两个子代细胞中去的。相反地，似乎大核的多倍体状态（四小核草履虫中 800n，嗜热四膜虫中 45n）在无性繁殖的过程中可以避免致死基因的丢失。大多的物种有一定的无性生活年限，克隆的细胞系如果不参与有性的繁殖，就会在它们衰老前死亡。

## 2. 接合：相互受精揭示非孟德尔遗传

纤毛动物是雌雄同体的物种，可以进行接合行为，接合是一种包括两个亲本细胞相互受精（reciprocal fertilization）的配偶过程。对于合适的具有一定无性史年龄的成熟细胞而言，轻微的饥饿会激活有性生活，不相同的细胞可配对起始接合作用。如果没有相容的配偶，某些物种就会进行自交，称为自体受精（autogamy）。在上述两种情况中，都会保证核重组过程的发生，从小核的减数分裂开始。细胞核行为的顺序在所有的物种中是相似，当然会有一些不同，图 7-1 中就阐述了四小核草履虫与嗜热四膜虫的核行为（Sonnerborn 1975）。

减数分裂后的发育起始于从每个细胞中挑选出一个单倍体核用于基因组信息的传递。挑出的细胞核进行一个额外的分裂产生两个遗传组成上相同的配子核。在接合中，

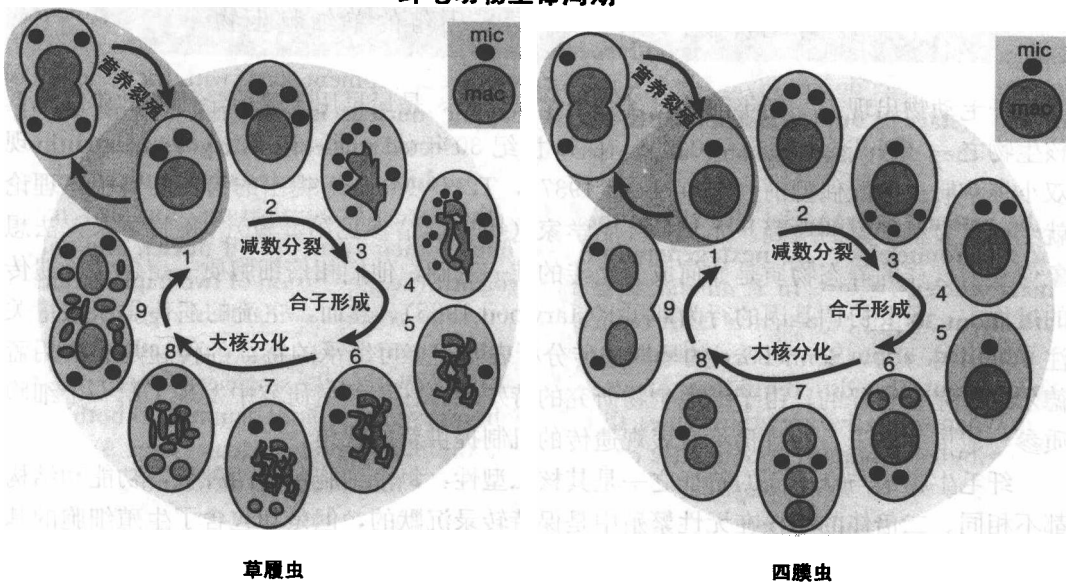



图 7-1 草履虫（左）和四膜虫（右）的生命周期

（发育第 1 阶段）营养细胞通过二分裂方式繁殖：在阶段 2～8 有性发育将在两个细胞接合或自核交配后开始（仅限于草履虫）。（发育第 2～3 阶段）小核减数分裂末期选择单倍体产物之一作为配子核而其他残留物将被降解。草履虫亲本大核开始呈分叶状。（发育第 4～6 阶段）合子形成：被选择的核再一次分裂后产生两个遗传上同样的单倍体核。接合期间两个同样的配子核之一在两配偶间交换，并且接下来配子融合产生二倍体合子核（红色）。自交期间，两个同样配子核简单地融合在一起。两次附加的合子后分裂产生未分化的小核和大核。（发育第 6～8 阶段）核分化：第二次合子后分裂后生成核中的两个成为新的小核，而另外两个开始分化成为大核（粉色）。在草履虫，母本的大核是碎片状的，而在四膜虫大核呈纹状。并且四膜虫的新的小核之一被降解。（发育第 9 阶段）核分裂：最初的无性分裂是特别的，因为新的大核无需分裂成为子细胞而小核通过有丝分裂分离到子代。最终，草履虫亲本大核的成分未分裂，而其余的在次第的分裂间随机分配直至丢失。

配偶的双方交换它们两个单倍体核中的一个，并产生合子（即两个单倍体核的融合）并在接合双方的体内都产生了遗传上相同的合子核（发育第 3～5 阶段，图 7-1）。在自体受精中，一个细胞中产生的两个合子核再融合产生一个二倍同源基因组的整体。在两种情况中，产生的双倍基因组体（第五阶段）分裂两次以上，四个子细胞分化命运不同：两个变成新的小核另两个变成新的大核（第 6、7 阶段）。核发育完成以后，会产生两种情况，第一个是两个新的小核都在新的无性系中保留，就如同四小核草履虫一样；第二个则是两个中的一个降解，就像在嗜热四膜虫那样。在两种物种中，两个新的大核在第一次细胞分裂时不分裂，而是均匀地分配到两个子细胞中（发育第 9 阶段）；它们在第二次无性分裂时才开始分裂。

当亲本的小核产生下一代新的大核和小核时，亲本的大核却丢失了。在四小核草履虫中，大核断裂成为 30 个片段以快速阻止 DNA 复制的发生，尽管在新大核分化的整个过程中，原来大核的转录依然活跃。当无性生活时重新开始时，那些碎片会毫无保留地随机分配到子代的细胞中。（发育第 9 阶段）在嗜热四膜虫中，亲本的大核不会断裂，

但是会发生固缩然后在第一次无性分裂前降解，其机制类似细胞凋亡（David et al. 1992）（发育第 7、8 阶段）。

## 3. 纤毛动物的胞质遗传

双核小草履虫的接合行为的生物学很适宜观察与细胞质效应相关的表观遗传现象。当相互受精产生遗传上相同的合子核时，细胞质部分几乎没有发生什么交换，于是可以很有效地将核基因的行为与环境的因素也就是母细胞的细胞质区分开来（图 7-2）。每次遗传杂交等同于对不同母亲产生的具有相同基因组的“双胞胎”的研究。Sonneborn 的研究表明两个亲本细胞之间不同的表型可以在其对应的后代中得以维持，即使这两个后代具有相同的基因型。在一些例子中，表型的区别后来被发现取决于核外基因也就说不能归入表观遗传现象的范畴。例如，双核小草履虫有一类 killer 品系，具有致死的性质是因为栖息在胞质中的红细菌种生物与其形成内共生关系，红细菌会释放一种毒素杀死敏感的品系（综述见 Preer et al. 1974；Pond et al. 1989）。其他的例子，例如血清型的母系遗传，血清型决定于几种平行同源的表面抗原基因中的一种的表达，是胞质因素对基因活性产生影响的很好的例子。

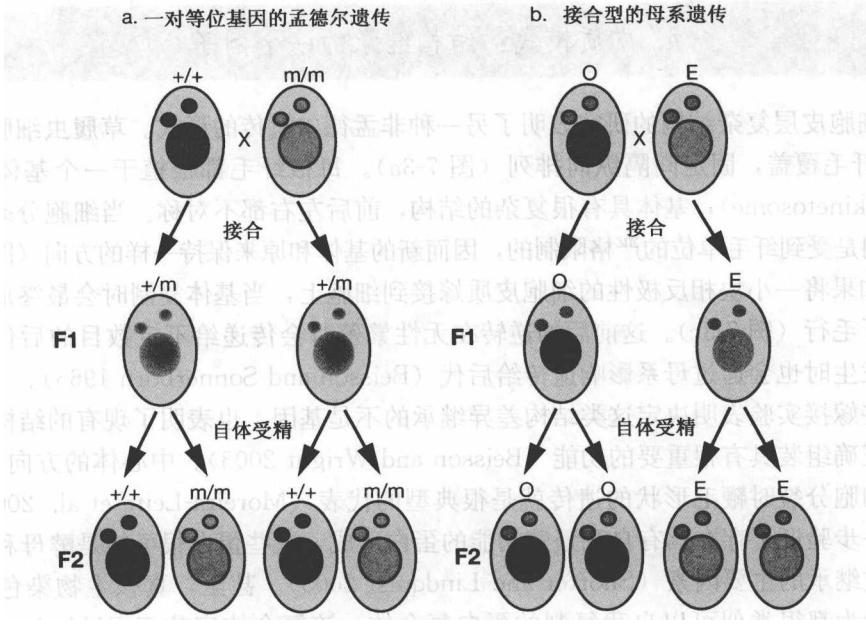



图 7-2　孟德尔规律对胞质遗传

（a）接合和自体受精揭示了两个草履虫细胞间的交配，其中每一个纯合子有不同的等位基因 $M$ 或 $m$。接合过程涉及两个样同配子核之一的相互交换。$F_1$ 有着相同的杂合子基因型个体。自体受精是一个自身受孕的过程，能在仅为一次有性传代的过程中产生完整的纯合子基因型，例如 $F_2$ 的这些个体有 50% 的机会成为 $M/M$ 或 $m/m$ 基因型。（b）$F_1$ 克隆间的表型差异揭示了胞质遗传特征。草履虫配偶类型（O 或 E）在全能生殖细胞小核（小环）转化而来的体细胞大核（大环）发育期间被不可逆性地决定；然而，亲本大核指导每个接合后个体朝向维持现存接合型分化。

如同血清型一样，四核小草履虫另两种配偶型，称为O型和E型，这两种配偶型便体现出胞质模式的继承特性。O和E的形状是终极分化的表型，决定于从全能干细胞衍生出的体细胞大核的发育。在接合作用以后，从O型亲本无性产生的后代几乎总是与O型配对，而从E型亲本无性产生的后代总是与E型配对，尽管两种接合后个体是从完全相同的合子基因组发育而来（图7-2b）。另外，当两个接合细胞间形成较大的胞质桥，胞质之间就可以进行很大规模的交换，那么两个亲本的后代通常都会发育成E型（Sonneborn 1977）。因此，肯定存在一种胞质因子指导E型的发育。

就像下面介绍的那样，生殖细胞基因组在大核发育阶段进行大规模的DNA重组，尽管现在没有确认决定配偶型的基因，但是一种影响配偶型决定的孟德尔式突变体被验证会在其他的位点干扰基因组的重排（Mayer and Keller 1996）。如果受调控的基因组重排可以决定配偶型，那么决定E型的胞质因子应该可以指导配偶型决定基因的另一种重排，使得该基因在大核中的形式是E型所特有的。在接合作用发生时，这种形式也势必产生决定E的胞质因子，这种因子是随后的继承性所必需的。就像在第10节中提到的，不同的重排模式可以从母系的大核遗传给合子的大核，这似乎是一种由RNA分子介导的反式作用的核效应，是同源依赖性的。因此，决定非孟德尔遗传的胞质因子似乎是一种RNA分子，它可以调控配偶型决定基因的发育“突变”。

### 4. 皮质模式：结构继承的一个例子

对细胞皮层复杂结构的研究表明了另一种非孟德尔遗传的形式。草履虫细胞由大约4000条纤毛覆盖，固定间隔纵向排列（图7-3a）。每根纤毛都根植于一个基体（basal body or kinetosome），基体具有很复杂的结构，前后左右都不对称。当细胞分裂时，基体的复制是受到纤毛单位的严格限制的，因而新的基体和原来保持一样的方向（图7-3b）。然而，如果将一小块相反极性的细胞皮质嫁接到细胞上，当基体复制时会最终形成完全逆转的纤毛行（图7-3c）。这前后的逆转在无性繁殖时会传递给不定数目的后代，在接合作用发生时也会通过母系影响遗传给后代（Beisson and Sonnerborn 1965）。

这些嫁接实验表明决定这类结构差异继承的不是基因，也表明了现有的结构对于新结构的正确组装具有很重要的功能（Beisson and Wright 2003）。中心体的方向性复制、锥体虫细胞分裂时鞭毛形状的遗传就是很典型的代表（Moreira-Leite et al. 2001）。朊病毒进一步验证了存在具有自我繁殖功能的蛋白形成，这些蛋白很可能是酵母和哺乳动物中胞质继承的主要因素（Shorter and Lindquist 2005）。甚至，真核生物染色体的中心粒的行为都很类似可以自我复制的蛋白复合体，该复合体定位于DNA上，却不由DNA来决定（Cleveland et al. 2003；见第14章）。这些表观遗传现象告诉我们不是所有的细胞内结构都可以通过基因上的信息重新组装的。从更广泛的意义上说，DNA自身的复制是一种结构继承的例子，但是基因组显然不是唯一的分裂细胞需要复制的结构。因此，表观遗传性的结构继承可以被视为是生命基础的机制之一。

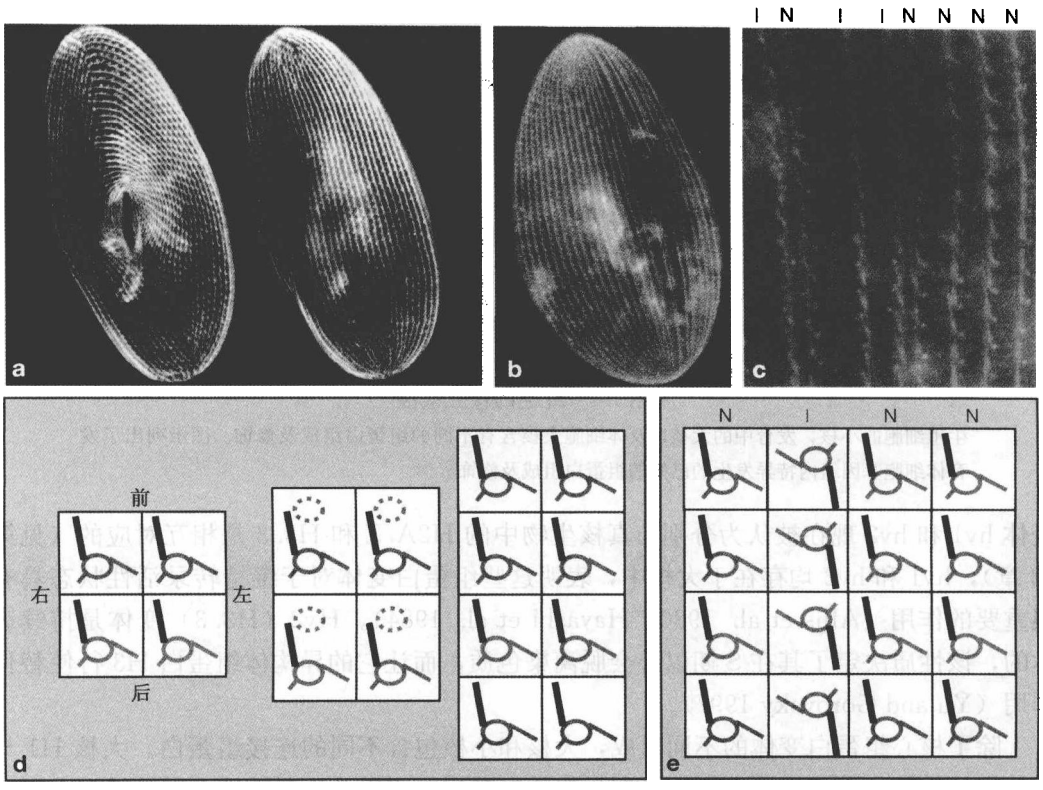



图 7-3　草履虫皮层单元极性的结构继承

(a) 基底小体和纤毛根的免疫标记揭示了野生型细胞皮层成行平行排列组织结构的规律。图示为腹面（左侧）和背面（右侧）观。(b) 细胞的背面观展示了由于少数皮层单元行的前后极性反转而造成的规律结构的破坏。基底小体免疫标记为红色，纤毛根为绿色。(c) 一片皮层的扩大图片显示纤毛根的反转方向，反转行 I 相对于对照行 N。(d) 基底小体（绿环）生长期间复制的示意图；右边为朝前方向的纤毛小根（紫色）和两个微小管环。复制以固定的几何排列形状发生：即每个新的基底小体被放置于亲本的前方以确保相同的极性。(e) 每行基底小体的重复复制不确定地维持均匀的走向（照片经 Janine Beisson 许可）。

## 5. 大核与小核：活性染色质对沉默染色质的模型

表观遗传学中一个基本概念是个体中一段 DNA 序列的不同拷贝可以呈现不同的活性，并且这种不同的状态可以得到稳定保持。纤毛生物的核二型性就是自然界中存在于同一个细胞质中的同源序列却呈现相反活性状态的例证。大核是转录活跃状态的代表，而小核则代表了抑制或沉默状态（图 7-4）。早期的生化和免疫组化研究，主要是在四膜虫中致力于比较这些不同核在无性细胞和有性发育阶段的不同性质，希望能够发现是什么决定了这些不同的活性状态，尤其是在染色质结构的水平。

### 5.1　大核与小核组蛋白的分离表明了不同组蛋白的不同功能

通过分离四膜虫中大核和小核的组蛋白，导致了特异的组蛋白变体的发现。组蛋白

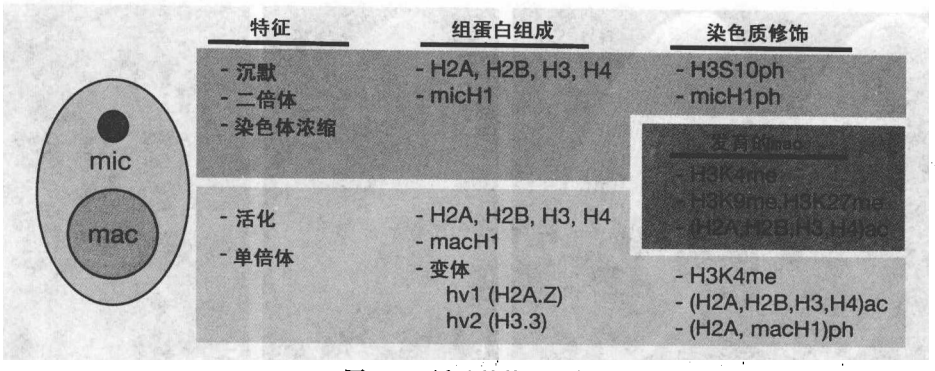



图 7-4 纤毛的核二型性

生殖细胞的小核，发育中的大核以及体细胞大核含有不同的组蛋白组成及修饰。图示列出了发育体细胞基因组内特异发生的已知的组蛋白组成及修饰。

变体 hv1 和 hv2 现在被认为分别与真核生物中的 H2A. Z 和 H3. 3 是相互对应的（见第 13 章），hv1 和 hv2 均存在于大核中，表明这些组蛋白变体对于维持转录活性状态具有很重要的作用（Allis et al. 1980；Hayashi et al. 1984）。Hv2（H3. 3）变体是持续表达的，该性质决定了其在 S 期以外会脱离染色质，而让它的异构体组蛋白 H3 行使替代作用（Yu and Gorovsky 1997）。

除了核心组蛋白变体的不同组成，大核和小核包含不同的连接组蛋白。大核 H1 与其他真核生物的连接组蛋白相比具有相似的氨基酸组成以及生化特性，但是缺乏中央球形结构域（Wu et al. 1986；Hayashi et al. 1987）。两种连接组蛋白的基因对细胞生存均非必需的。有趣的是，将任何一种连接组蛋白敲除，会导致其相应的核的体积变大，提示连接组蛋白应该在染色体压缩上具有功能，更可能是稳定高级结构的染色质状态（Shen et al. 1995）。缺失大核的 H1 样蛋白同时会导致特定基因的表达变化，表明该连接组蛋白也参与了适当的转录调控的维持（Shen and Gorovsky 1996）。

## 5.2 染色质修饰与染色质的活性状态相关

### 乙酰化

大核中组蛋白的高乙酰化和小核中乙酰化修饰的缺失提供证据表明这种转录后的修饰与基因活性有重要的关联（Vavra et al. 1982）。20 世纪 90 年代中期以前，人们没有发现任何生物的组蛋白乙酰化酶，直到 20 世纪 90 年代，C. David Allis 和他的同事纯化了第一个 A 型（核内的）组蛋白乙酰化酶（histone acetyltransferase，HAT）（Brownell and Allis 1995；Brownell et al. 1996）。这些研究者从高度提纯的大核入手，将这种乙酰化酶的活性与 B 型胞质的 HAT 活性分开，并实现使用胶内系统纯化。在这过程中，纯化的组蛋白聚合在聚丙烯酰胺介质中用于蛋白组分的分离。电泳以后，蛋白复性并与带放射标记的乙酰辅酶 A（acetyl-CoA）一起温育，发现一条可以将标记整合入组蛋白基质的、明显的分子量为 55 KD 大小的蛋白链。

真正的突破产生于对纯化蛋白的测序以及基因克隆。这个四膜虫的 HAT 被发现与面包酵母中一个明确功能的转录因子 Gcn5 蛋白同源。在该发现之前，转录激活子被认为主要通过将 RNA 聚合酶募集到启动子上来工作，但是这项工作却表明转录激活子也可以具有酶的活性，修饰染色质或者其他的转录调控因子，因而改变目的基因的状态。由此，潘多拉盒子被打开，很多已知的调控因子被证明是 HAT。

甲基化

纤毛动物的核二型性又一次证明它们在揭示组蛋白甲基化功能上的优势。这种修饰仅限于生长的四膜虫细胞的大核中（图 7-4）。从这些核中纯化得到的组蛋白赖氨酸甲基化转移酶（histone lysine methyl-transferase，HKMT）的活性特异地指向 H3 的第 4 位赖氨酸（H3K4me），为这种特定的修饰和转录活性间的联系提供了早期证据（Strahl et al. 1999）。

组蛋白 H3 的第 9 位赖氨酸甲基化不存在于无性生长的细胞中，但是特异发生于大核发育时某段特定的生殖细胞序列上，而这段序列是不存在于体细胞基因组的（Taverna et al. 2002）。在某些特定序列上 H3K9me2（双甲基化）是受发育调控的，该现象为解释这种修饰对于同等意义上的异染色质位点提供了很好的模型（详细描述见下文）。

磷酸化

从大核、小核中纯化得到的带$^{32}$P 发射性标记的组蛋白表明 H3、H2A 和连接组蛋白是高度磷酸化的（Allis and Gorovsky 1981）。大核中 H1 很多位点都高度磷酸化，而且该修饰被证明参与特定基因转录的调控。使用突变体分析，Dou 和 Gorovsky 发现这种磷酸化的状态可以通过在 H1 上增加额外带电的氨基酸来模仿（Dou and Gorovsky 2000、2002）。带电的氨基酸无需存在于磷酸化的氨基酸对应的位置上，但是同等的效应却需要大量的带电位点。这些研究表明每个丝氨酸的磷酸化是不需要的，但是一定量的电荷密度可以增强特定的转录。

第 10 位丝氨酸这个单一位点，在组蛋白 H3 上的被磷酸化（H3S10ph）（Wei et al. 1998）。该修饰是依赖于细胞周期的，在很多真核生物中都与有丝分裂相关联。在四膜虫中，它仅限于小核的有丝分裂和减数分裂中。用 Ser10 号（S10A）位置被丙氨酸替代的突变体替代正常的 H3 基因会导致小核分裂的缺陷，造成染色体松弛以及染色体数目紊乱（Wei et al. 1999）。大核无丝分裂却不受影响。这些结果表明 H3 的磷酸化在染色体的压缩/分离都有重要的作用。纤毛动物独特的核二型性又一次为染色质的修饰提供了关键的注意点。

## 6. 整个基因组范围的重排发生于大核发育时期

四核草履虫与嗜热四膜虫体细胞基因组（大核基因组）的序列显示了很高的基因数量（分别为约 40 000 和约 27 000），但是基因组很小（分别为约 72Mb 和 104Mb）。这种基因组的结构有利于高效的基因表达（图 3-15，其他真核生物间的相互比较）。体细

胞基因组的这种精简结构是由从生殖细胞来源的染色质的大量重排形成的，重排发生于大核发育的特定阶段（综述见 Prescott 1994；Coyne et al. 1996；Klobutcher and Herrick 1997；Jahn and Klobutcher 2002）。因此，生殖和体细胞基因组的核差异不仅体现在刚刚提及的建立不同的染色质阶段，还包括染色体结构上的重新组织（physical reorganization）。

通常观察到的有两种重排形式，是研究使用的所有纤毛动物所特有的，内源 DNA 缺失和染色体片段化（图 7-5）。在其他真核动物中这种现象也存在，例如哺乳动物淋巴细胞的 VDJ 位点的重组（见第 21 章）。就纤毛动物这种物种而言，重组事件会从每个新产生的大核中移除 10%～95%的生殖细胞基因组。基本上所有的重复序列，包括转座元件都被移除了，当然这在一定程度上是因为大核基因的高密度。重排位点的数目在不同物种中也不同，从四膜虫的 6000 个到腹毛目纤毛虫的大约 100 000 个。整个基因组层面上这种高度调控和可重复事件的发生为我们了解细胞如何识别和作用于特异 DNA 片段提供了研究的机会。对于纤毛动物 DNA 重组的研究为表观遗传机制提供了独特的切入视角，尤其是与同源依赖性过程相关时。下面对于 DNA 重组事件的描述将为后来表观遗传调控的讨论提供足够的背景知识。

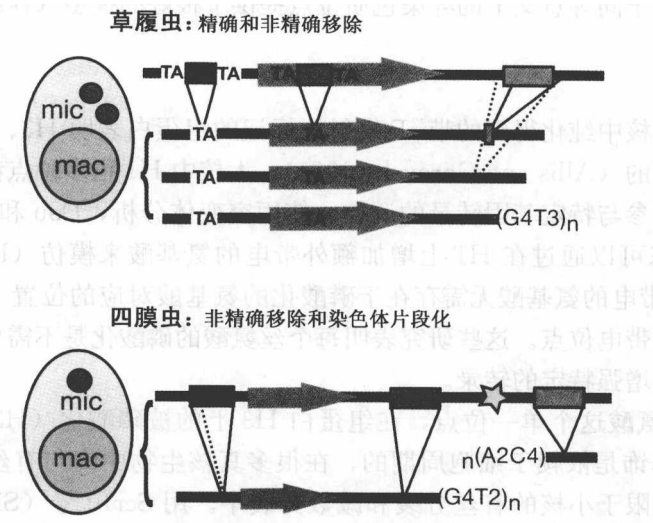



图 7-5 纤毛的 DNA 重排

新的大核发育期间会发生染色质片段延伸及 DNA 消除。（上图）草履虫的 DNA 重排包括 TA-结合 IES 的精确消除（编码区［箭头］和非编码区的有色条）以及非精确消除（橘黄色条），后者导致改变的 DNA 消除和片段化（G4T3)$_n$；（下图）四膜虫大约有 6000 个基因位点发生 IES（有色条）的精确消除，染色质片段化是一个 15bp 的保守序列，即染色体片段化序列（CBS）（星形）。

## 6.1 内源 DNA 缺失：精确和非精确机制

### 精确缺失

精确缺失（基因内内源移除序列）发生于大核染色体所有拷贝的同一个核苷酸位

点。内源移除序列（internal eliminated sequences，IES）是短小、单拷贝的 DNA 片段，通常是从编码区域移除，却也在生殖细胞染色体的基因间和内含子区域发现过（图 7-5）。精确切除的 IES 两端都连接有短重复序列，重复序列在不同的物种中不同。有突出的一类被称为“TA”的 IES，是在草履虫和某些旋毛纤毛虫例如厚游仆虫中发现的，识别方法是：它们连接区的 5′-TA-3′重复序列是不变的，它的一个拷贝在大核的一些位点切除之后仍旧保留（Betermier 2004）。TA 碱基之间的核酸序列并不是随机的，而是形成了一种简并序列，与 Tc1l/mariner 转座子的末端很相似。因此，这些 IES 可能由早期这类移动元件的插入进化衍生而来（Klobutcher and Herrick 1997）。尽管如此，很多 IES 的末端与进化保守序列（consensus sequence）的相符性很差，而在非切除的区域却发现了与进化保守序列十分吻合的序列，这表明进化保守序列本身并不能提供足够的信息来指导对四小核草履虫每个单倍体基因组约 60000 个 IES 的特异剪切，因而使得它们如何识别成为一个很重要的问题。

一种不同于精确缺失、令人难以想象的过程发生在旋毛纤毛虫的亚类中，即 IES 被移除的同时基因的顺序却没有打乱（Prescott 1999）。在小核中，基因是被打乱的，具有大核命运的序列（macronuclear destined sequences，MDS，即 DNA 的“外显子”）不仅被 IES 隔开，而且其顺序相比重排以后大核的线性排列而言是杂乱的。在生殖细胞中，两个 MDS 合并在一起可以组成表达基因的部分可能相隔很远，有的时候甚至在非连锁的位点（Landweber et al. 2000），甚至可能两部分的排列是相反的方向。这种重排的精确性似乎是被相对的同源长重复序列指导（平均 11bp），与同类的 MDS 末端具有相同序列，但与其他的 MDS 末端没有关联。尽管这些长重复序列归功于基因排列的精确、有序，但是并没有证据显示它们是必须的，因而就引起了一种假设，有预先存在的模板（Prescott et al. 2003）。

非精确缺失

尽管 IES 从体细胞基因组中移除是高效且可重复的，而在一些情况下，在多倍体的基因组上独立的切割形成的边界是互不相同的（图 7-5）。这种不同体现在四膜虫的数十个碱基不同到草履虫的数千个碱基不同。非精确缺失（基因间重复序列）是所有用于研究的四膜虫 IES 的特点，也是移除类似基因间转座子或者草履虫中的小卫星等重复序列的主要方式（Le Mouël et al. 2003）。就像精确缺失一样，这些缺失主要发生在短小的重复序列之间，其中的一段在大核中有所保留。在四膜虫中，重复序列在已知的 IES 中都是不同的，而在草履虫中，重复序列总是包含至少一个 TA 核苷酸对，表明草履虫的基因缺失机制可能与精确缺失相关。

## 6.2　染色体片段化

在大核的发育阶段，由生殖细胞衍生而来的染色体会断裂成为小分子，之后两端连上端粒重复序列（图 7-5）。于是产生的大核染色体乍看如同缺失了中心粒的染色体，因此，染色体片段化可能有利于这些分子在大核无丝分裂过程中的平均分配。片段化的程度在不同的物种中都不相同，在旋毛纤毛虫中，这种现象到了极致，产生微小的“纳

米级染色体（nanochromosomes）”，这种染色体通常包含一个基因，而在草履虫和四膜虫中，大核基因组的大小从 20kb 到大于 1Mb 不等并包含很多基因。

在大多纤毛动物中，这个过程是不精确的，因而端粒重复序列的确切位置也是不同的，且经常导致生殖细胞序列的缺失。厚游仆虫是一个例外，因为它的端粒总是夹在同样的核苷酸位点（Klobutcher 1999）。在四膜虫中，存在一个保守的 15bp 的染色体断裂序列（chromosome breakage sequence，CBS），大约在生殖细胞染色体中发现了约 280 个位点（Cassudy-Hanley et al. 2005；Hamilton et al. 2005），CBS 对于指导片段化和新的端粒附加是必须且重要的（Fan and Yao 1996）。相反地，在小核草履虫物种中没有发现相似的 CBS。甚至，所有的证据表明染色体片段化式非精确缺失的另一个结果（图 7-5）是被移除 DNA 有两种命运：或者侧翼序列重新结合，或者加上端粒（Le Mouël et al. 2003）。

## 7. 基因组重排的机制

尽管纤毛动物基因组重排研究已经进行了 30 多年，这些事件的分子机制仍旧大量未知（Yao et al. 2002）。部分的挑战来自于解释为何不同的纤毛物种发生的重排事件具有很大的差异性。好几种纤毛动物的 DNA 切除中间体和环形的切割产物都有过描述（Jaraczewski and Jahn 1993；Klobutcher et al. 1993；Williams et al. 1993；Saveliev and Cox 1995、1996；Betermier et al. 2000；Gratias and Betermier 2003）。数据无法推导出一个统一的切割机制，即便是在 IES 末端进化保守序列非常相似的纤毛动物物种间也是如此（Klobutcher and Herrick 1995；Gratias and Betermier 2001；Betermier 2004）。

尽管切割的机制多种多样，这些指导 DNA 重排的事件总会存在一些重要的共有特征。一个相似点就是类似转座子和重复序列更容易被移除。就像在这本书其他部分提到的，表观遗传机制的一个作用就是抑制这些潜在可能被移除的 DNA 元件的活性。允许转座子从表达沉默的生殖基因组跳到高度活跃的大核基因组会对体细胞基因组造成灾难性后果。很多证据表明了与 RNA 干扰和异染色质形成相关的机制参与 DNA 重排的过程（Mochizuki and Gorovsky 2004；Yao and Chao 2005）。

在大多数真核生物中，在 H3 组蛋白第 9 位赖氨酸甲基化（H3K9me）与抑制表达的 DNA 是很相关的，这些 DNA 分块地位于核异染色质的区域（见第 3 章第 7 节）。在四膜虫中，这种修饰在转录沉默的小核中并没有发现，但是在发育的大核以及同时发生的 DNA 重排中都有发现（Taverna et al. 2002）。染色质免疫沉淀试验表明这种修饰在于 IES 相关的组蛋白上频繁出现。DNA 移除和异染色质的形成最初被认为与包含染色质域的 DNA 程序性降解蛋白 I（programmed DNA degradation 1，简称 Pdd1）相关，Pdd1 蛋白在发育时期大量表达，与包含将从体细胞基因组移除的生殖细胞 DNA（即 IES）的位点共定位（Madireddi et al. 1994、1996）。Chromo 结构域是对特定的甲基化组蛋白残基有亲和力的蛋白基序（protein motif）。也许染色质调控中染色质域蛋白介入的原形模型是在果蝇细胞中异染色质蛋白 1（heterochromatin protein 1，HP1）结合

甲基化的 H3K9，同时参与异染色质形成（更多细节请见第 5 章）。Pdd1 和 Pdd3 的 Chromo 结构域会结合到 H3K9 双甲基化的肽链上，Pdd1 和 Pdd3 是 DNA 重排必需的两个蛋白。为了证实这种染色质的修饰是 DNA 重排所必需的，Liu、Mochizuki 和 Gorovsky 使用 K9Q 赖氨酸 9 位突变体替换主要 H3 基因。这些细胞无法在发育阶段有效地移除 IES 片段，但是 Pdd1 却很好地定位在大核的前体中，因而表明 H3K9 的双甲基化是 DNA 移除所必需的。因为染色质状态的形成是表观遗传调控的重要决定因素，那么很重要的一个问题是，H3K9 双甲基化的标记如何特异性针对即将被移除的 DNA 片段？如下面所描述的，几个实验证实同源 RNA 参与指导了 DNA 的重排。

## 8. 纤毛动物中同源依赖性的基因沉默

同源依赖的、RNA 介导的基因沉默机制广泛存在于真核生物的表观遗传调控中。这类机制同时存在于纤毛动物的证据首先在草履虫中发现将非表达转基因转入无性大核后可以造成与孟德尔遗传突变相同的表型。相似的作用也存在于草履虫和旋毛纤毛动物中，通过给细胞补充双链 RNA 而实现沉默，由此表明 RNA 干扰通路参与其中，结果这些通路中的一条引起了基因沉默导致 DNA 移除。

### 8.1　转基因诱导的基因沉默

草履虫的大核很容易通过显微注射转化入外源基因，因为任何引入的 DNA 片段都可以以高拷贝数存在于基因组中，自行复制不需要任何起始。转入高拷贝数的非表达转基因可以引起序列高度相似的内源基因的转录后沉默（Ruiz et al. 1998；Galvani and Sperling 2001）。如果该基因的 3′UTR 区域存在，那么沉默就无法实现，表明 RNA 中存在的调控信号会影响沉默的能力（Galvani and Sperling 2001）。结果，发现沉默现象还与同源性的短小的、大约 23 nt 长的 RNA 的累积有关联（Garnier et al. 2004），表明 RNA 干扰通路参与其中。那 23 nt 的小 RNA 似乎对同源 mRNA 的特异降解有作用，因而被称为 siRNA（short interfering RNA）。一类相似的 23～24nt 的 RNA 被发现存在于无性世代的四膜虫细胞中，虽然没有数据表明其可能的功能，但似乎它们代表的就是内源的 siRNA（Lee and Collins 2006）。

### 8.2　基因沉默是由双链 RNA 诱导的

双链 RNA（dsRNA）似乎是在草履虫引起转基因诱导的基因沉默的主要原因。转基因的沉默效率决定于某种异常 RNA 分子产生的量，这种 RNA 分子会对注入序列的正义和反义链产生反应。另外，双链 RNA 行使基因沉默的能力可以通过给草履虫细胞喂食表达双链 RNA 的大肠杆菌来验证，这种方法原本是从线虫上衍生而来的（Timmons and Fire 1998；Timmons et al. 2001）。内源基因的沉默短到在三代无性分裂后即可以看到表型，就是说小于 24 小时内（Galvani and Sperling 2002）。如果对旋毛纤毛生物喂食热击致死的大肠杆菌可以达到同样的基因沉默的目的（Paschka et al. 2003），而旋毛纤毛生物通常是喂食藻类的，这现象表明很大一类的纤毛生物具有这种机制。在

草履虫中，分子生物学分析表明喂食双链 RNA 会导致 23 nt 的 siRNA 的累积，这种累积现象也在转基因诱导的基因沉默中发现（Nowacki et al. 2005），表明这两种现象依赖共同的 RNA 干扰通路。

通过对草履虫喂食双链 RNA 诱导的基因沉默可以通过将食物替换成常规的细菌而恢复；相似的是，将双链 RNA 直接注入细胞质内只会诱导瞬时的同源基因沉默，推测是因为注射的双链 RNA 在无性生殖的过程中被迅速的稀释（Galvani and Sperling 2002）。这些研究表明双链 RNA 不可能在营养细胞的细胞质中自行扩增到很高的程度，然而线虫体内的双链 RNA 却有既定的命运，可以导致基因沉默成为可以继承的状态。这进一步表明草履虫的 RNA 干扰不会引起大核中转录基因的稳定沉默。可继承的沉默似乎需要组蛋白 H3K9 甲基化，就像上面所提到的，这种修饰在营养细胞的大核中是缺失的，至少在嗜热四膜虫中是这样的。

## 9. 基因组重排是同源依赖机制指导的

在纤毛动物的发育过程中，三个不同的基因组会区分开并进入不同的发育命运：生殖细胞的小核基因组在体细胞中新的大核基因组大规模进行重排时保持不变，而母体的体细胞基因组则注定被破坏。在这个更广泛的框架中，纤毛动物研究人员致力于理解 DNA 重排模式的重复性。最初的努力试图发现顺式作用的、可以募集重组蛋白的 DNA 序列，但是这种搜索几乎没有成果（Yao et al. 2002；Betermier 2004）。初始 DNA 序列显然不是指导大核基因重组的唯一决定因素，这个结论有个很有力的证据：H3K9 甲基化标记出现在被移除的 DNA 片段上。另外，很多重排模式对同源性依赖作用是很敏感的，就像下面描述的那样，允许不同的模式通过母体继承给有性后代，并不符合孟德尔关于野生型生殖细胞基因组遗传的理论。因此，重排模式的继承符合表观遗传现象的定义。

### 9.1 实验诱导大核发育中的特异性缺失

在草履虫中的转录后基因沉默现象被报道以前，科学家就已经发现将高拷贝数的克隆片段转入营养细胞大核会改变其有性后代的 DNA 重排。更令人称奇的是，与转基因同源的序列会通过非精确的机制被特异性的移除，而小核基因组没有任何变化（图 7-6）（Meyer 1992）。该现象具备一定普遍性，因为所有试验的 DNA 片段都会导致 DNA 的移除（Meyer et al. 1997）。这些实验表明序列特异性的信息在有性生殖阶段通过细胞质从母体大核传递到发育中的合子大核。这种影响的普遍性否认了一种说法，即是注入转基因编码的序列特异性结合蛋白的作用。符合这种观察结果的最保守的解释认为这是核酸的作用，可能是 RNA 分子，在核间传递信息，通过配对作用识别它们的目标。

一些能够有效引起合子后阶段 DNA 移除的是非表达转基因，例如包含移码突变或者切除了 5′或 3′UTR 的片段（Garnier et al. 2004）。那些可以产生稳定的、可翻译的 mRNA 不会促进同源基因从发育的大核中移除。同样，能促进 DNA 移除的转基因也可以在被转入克隆的营养性生长阶段引起内源母体基因的沉默。因此，合子后阶段的

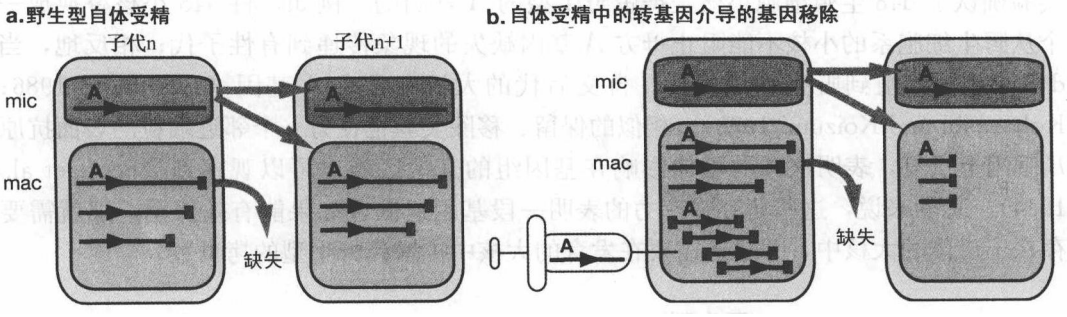



图 7-6　草履虫的转基因诱导的 DNA 消除

(a) 野生型 *A* 基因（红色箭头）位于靠近大核染色质异源基因的末端；粉色框为大核端粒。小核染色质的重排模式忠实地从一代复制到下一代。(b) 当新的大核从野生型生殖细胞发生时，多拷贝 *A* 基因引入母本大核诱导了有性后代内源 *A* 基因的完全消除，新大核的端粒恰好位于 *A* 基因的上游。

DNA 移除与合子前阶段的基因沉默相关，也与 23 nt 的 siRNA 的积累有关。siRNA 的水平后来被证实在细胞整个发育阶段都维持在一定水平，表明它们可能与这些移除的引发相关。

如果协助基因沉默的 siRNA 也指导 DNA 重排，那么在发育前或者发育中引入双链 RNA 都会引起递减的后代中同源序列的移除。为了验证这一点，科学家给草履虫喂食产生 ND7 编码 DNA 同源的双链 RNA 的大肠杆菌，nd7 是参与调解称为刺丝泡的分泌小泡的胞外分泌的基因，这种草履虫进行自交并从野生型的小核中发育出新的大核。很大数量的自交后代呈现出 ND7 突变表型，因为该基因已从大核中移除（Garnier et al. 2004）。甚至表型上仍为野生型的后代也显示出了 ND7 拷贝部分缺失。就像那些与转基因诱导的基因沉默相关的 siRNA 一样，与双链 RNA 相关的 23 nt 的 siRNA 也是在整个自交过程中细胞中都有分布，确认了它们与合子后阶段的基因移除相关（Nowacki et al. 2005）。由双链 RNA 指导的 DNA 移除并不仅限于草履虫，对接合四膜虫显微注射体外转录的正义和反义 RNA 会导致这些序列不精确的移除，这些 RNA 是与大核中某些基因位点同源的（Yao et al. 2003）。

## 10. 重排模式是由生殖细胞和体细胞基因组的比较而决定的

### 10.1　d48 为例：其他重排方式的表观遗传

尽管转基因和双链 RNA 诱导的重排表明了 DNA 移除过程的表观遗传本质，更令人惊奇的是发现这种诱导的重排模式在经过后续的大核发育中是可继承的。重排的表观遗传调控的第一个证据是发现四核草履虫编码 *A* 表面抗原基因的缺失在与野生型交配后是可以通过胞质继承的（图 7-7）（Epstein and Forney 1984）。在野生型的大核中，*A* 基因位于一条染色体的末端，距离端粒 8、13 或者 26kb，决定于三种不同的 DNA 片段化位点。一种细胞系 d48 被发现缺乏 *A* 基因表达功能因为基因本身在整个下游序列中都是缺失的，因为它的端粒在基因的 5′端形成（Forney and Blackburn 1988）。核移植

实验确认了 d48 生殖细胞小核是携带野生型的 *A* 基因的。例如，将 d48 小核替换成一个从野生细胞系的小核不能阻止母方 *A* 基因缺失的现象传递到有性子代；相反地，当 d48 小核被移植到野生型细胞后，自交后代的大核中含有 *A* 基因（Harumoto 1986；Kobayashi and Koizumi 1990）。相似的保留、移除实验也在另一个邻近端粒的表面抗原 *B* 基因上重复，表明这样的母体影响在基因组的其他区域也可以观察到（Scott et al. 1944）。总的来说，这些研究强有力的表明一段基因组区域如果能有效保留，那就需要存在于母体的大核中，并且还需要在发育的大核中扩增到野生型的拷贝数。

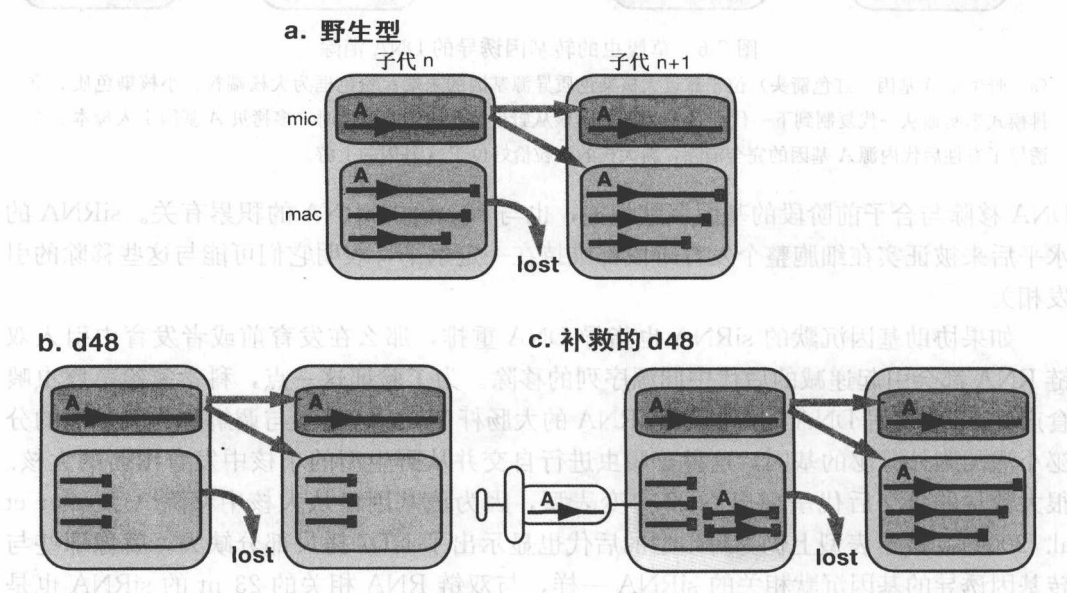



图 7-7 大核 *A* 基因消除的表观遗传和实验补救

(a) 野生型细胞系。(b) d48 细胞系，其大核里缺少 *A* 基因，但有野生型小核。*A* 基因在大核发育过程每次传代复制时被移除。(c) 在有性生殖大核发育过程中，d48 系用含 *A* 基因序列的大核内转化会特异地重建生殖细胞 *A* 基因的扩增。

## 10.2 实验诱导 DNA 缺失产生的表观遗传

正如在 d48 中所观察到的，母体对 DNA 重排的影响控制着很多甚至所有草履虫大核基因组的区域。甚至，由实验造成的有性后代大核的 DNA 移除可以自发的传递到更下一代中，通过母系继承的作用模式。这在高拷贝数转基因——*G* 基因引起的基因移除的草履虫中观察到，*G* 基因是一种亚端粒表面抗原基因（Meyer 1992），这种现象同时在由喂食双链 RNA 引起的 ND7 基因移除的草履虫中观察到（Garnier et al. 2004）。在任意一种情况中，诱导的转基因或者外源引入的双链 RNA 在后代的基因移除中不再需要。在 d48 和这些诱导改变的细胞系中，体细胞基因组的状态在自交后偶尔能够回复到野生型重排的模式，这进一步确认说基因本身是存在于小核基因组的，更突出的这种遗传使用的是表观遗传的模式。

在大核发育时期使用转座子和 IES 产生的基因沉默以及导致的发育 DNA 的移除在

后代细胞中是可“记忆”这点是很重要的。每个有性世代的重复性的基因移除似乎不是 23 nt 的 siRNA 引起，因为 siRNA 只在实验诱导的第一代细胞中。另外，对这种细胞系的遗传分析表明小核基因不会携带任何永久的印记，因为当它与野生型的大核一起通过接合作用转移到一个细胞系中时，通过大核的发育很容易就可以扩增。可见，简单地说该基因在发育的大核中是被移除了，因为母体大核中它就是缺失的。

## 10.3　引入小核基因组的外源序列的“自发性”移除

母体重排状态的可遗传性表明即将被重排生殖细胞基因组与现存的已经重排的基因组是经过比较的，在上一代基因组中缺失的序列会成为下一代形成的体细胞基因组被移除的对象。如果真的如此，那么引入生殖细胞小核的转基因就可以引起新的大核的基因移除。在四膜虫种，小核转化已经是可行的，研究人员发现整合入的用于基因干扰实验的药物抗性标记在连续几轮的接合作用是会从大核基因组中移除（Yao et al. 2003；Liu et al. 2005）。转基因的基因组位置，还有引入生殖细胞基因组的转基因拷贝数都会很大程度上影响移除的效率（Liu et al. 2005），表明该种纤毛动物中这种现象并不是普遍诱导的。尽管如此，这些结果已经证实一段外源序列，细菌的 *neo* 基因，可以被细胞识别为一段 IES。

## 10.4　可继承的大核 DNA 缺失的实验补救

母体体细胞大核中某段基因组区域的丢失是否足够导致其将来的移除呢？如果是这样的话，重新将 *A* 基因导入 d48 大核应该可以挽救 *A* 基因在发育过程中的丢失。最初显示在 d48 细胞发育早期将野生型的大核核质注射入 d48 的大核或自交的细胞之中（Harumoto 1986；Koizumi and Kobayashi 1989），都是导致 d48 在野生型细胞中的移除现象被永久性逆转。结果，将几种不重叠的 *A* 基因（覆盖达 8kb 的编码区）的直接注射入 d48 大核中正式 *A* 基因序列本身就已经足够回复野生型大核重排的模式（图 7-7）（Koizumi and Kobayashi 1989；Jessop-Murray et al. 1991；You et al. 1991）。

母体基因组对 DNA 重排的影响体现出了序列特异性。*A* 和 *B* 基因编码序列在 74% 水平上是相同的。尽管如此，将 *A* 基因的序列注射入某细胞系同时缺失 *A*、*B* 基因的大核中只能阻止 *A* 基因的移除；同样的，注射 *B* 基因序列也只能阻止 *B* 基因的移除（Scott et al. 1994）。另外，d48 大核的移除无法通过单细胞草履虫的 *G* 基因的转化来挽救，而 *G* 基因本身与 *A* 基因的相同性达到 78%。另一方面来说，*A* 基因的大核移除现象可以通过转化 *A* 基因的另一个等位基因来挽救，这两个基因具有 97% 的相似度（Forney et al. 1996）。因此，大核移除的挽救是一个依赖同源性的过程，它不需要任何特异的序列，但是要求一定程度的序列相似性。

d48 的母体移除作用的挽救表面上与转基因诱导的移除现象相违背，因为在这些实验中，将 *A* 基因序列注射入母体大核中恰恰导致了 *A* 基因完全相反的状态。这个明显的矛盾最终得以解决是因为发现合子后阶段的移除效应是依赖于转化克隆的依赖同源性基因沉默的（Garnier et al. 2004）。相反地，这种挽救效应也只有在不会引发基因沉默的转化情况下才会出现（Garnier et al. 2004）。因此，看起来 23 nt 的 siRNA 的积累会

阻止 *A* 基因序列在母体大核中的挽救效应，这会加速 *A* 基因在发育大核中 *A* 基因扩增。这表明介导反式核作用的细胞质因子是一个 *A* 基因的转录产物。然而，它没有必要是全长的 mRNA，因为即使编码序列的片段也具有挽救移除作用的活性。另外，包含整个基因的克隆往往在整个无性生殖的阶段表达 mRNA，相反胞质挽救因子的产生局限于核重排的那个阶段。

## 10.5 实验性抑制发育中大核的 IES 移除

### 草履虫中 IES 移除的同源依赖性抑制

如果细胞基因的移除或者维持是由体细胞和生殖细胞基因组成分比较决定的，那么甚至是 IES 的正常高效切割都会受到影响，如果在母体大核中存在拷贝的话。科学家发现孟德尔突变体 $mtF^{E}$ 在大核发育阶段具有多向性的作用，包括对于配偶类型的决定（Brygoo and keller 1981a、b），为上述预测提供了验证的机会。这种突变体，如果是纯合的，被发现会阻止 *G* 表面抗体基因的 IES 切除。令人惊奇的是，当 *mtF* 基因的野生型等位基因重新通过接合作用导入突变系列后，IES 在之后的大核发育过程中仍旧没有被切除（Meyer and Keller 1996）。对于这种变化细胞——称之为 $IES^{+}$ 的遗传分析确认了从遗传上说这就是野生型，IES 特异的保持力是在有性生殖过程中从母体那里继承而来的（Duharcourt et al. 1995）。

发育中的大核的 IES 的切除是由正确重排的 *G* 基因的母本拷贝诱导的，还是受到母本拷贝中的 IES 保留的 *G* 基因的抑制？为了回答这个问题，$IES^{+}$ 或者 $IES^{-}$ 细胞通过显微注射携带 *G* 基因序列的质粒的方式引入 *G* 基因（图 7-8）（Duharcourt et al. 1995）。在自交以后，转化的克隆里包含 $IES^{+}$ 质粒的会产生新生大核包含 IES 的后代，而 $IES^{+}$ 细胞转化 $IES^{-}$ 质粒则无法诱导移除，他们的后代则仍旧保持 $IES^{+}$ 的状态。一个只包含 IES 区域的质粒，没有任何侧翼序列的质粒，也会出现 IES 的保持，表示仅仅是母体的 IES 拷贝本身是抑制合子大核的移除的（Duharcourt et al. 1995）。

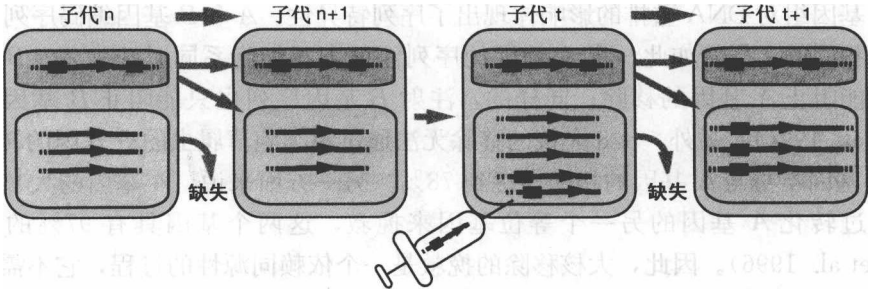



图 7-8 母本大核 IES 切除的同源依赖抑制

正常发育过程中，IES（红和绿条）被有效地切除。然而，携有既定 IES（初始转化为子代 t）的母本大核的转化能抑制次第子代（子代 t+1）（和以后子代）期间新大核分化的同源 IES 消除。

### 草履虫中“母本控制”以及“非母本控制”的 IES

其他 IES 的移除是否可以受到相似的母本效应的控制？这个问题通过将包含大段的小核 DNA（$IES^+$）的野生型细胞的大核转化得到回答，该段 DNA 包含 *G* 或者 *A* 表面抗原基因（Duharcourt et al. 1998）。注射的片段分别包含 6 个和 9 个 IES。实验发现 13 个 IES 片段被移除，而 5 个 IES 仍旧保持在自交的转化细胞的后代细胞的大核中。注射含单个 IES 的质粒的结果表明抑制作用是十分特异的：5 个 IES 中的任何一个仅仅诱导其同源合子 IES 片段的保持，但不会影响其他的 IES。对照的 DNA 片段则是包含大核（$IES^-$）*G* 基因，对任何 IES 都没有影响。在 13 个实验的草履虫 IES 中，没有明显的大小差异，碱基组成的差异或者在基因位置上的差异，因而 5 个非移除和 8 个移除 IES 并无明显区别（Duharcourt et al. 1998）。

### 四膜虫中同源依赖性的 IES 移除效应的抑制

与草履虫中相似的实验表明四膜虫大核的 DNA 成分可以调节发育中的体细胞基因组同源序列的移除。两个很有代表性的 IES，M 和 R 移除片段通过显微注射的方式注射入野生株的大核中，这样它们可以保持很高的拷贝数。当细胞被诱导进行接合作用时，包含母体拷贝 M 片段的后代无法有效地移除对应的 IES（Chalker and Yao 1996）。同样的，那些父系细胞被注射入 R 片段的也无法移除其同源的 IES，而对同源片段的移除作用的显著性抑制并没有发现。因此，DNA 移除的一直是序列特异性的。与 IES 同源的序列是足以引起抑制的，但是其侧翼 DNA 却没有效应。重要的是，这种 DNA 移除作用的丧失是可以继承的，也是说后代的基因组中仍含有这些 IES 的拷贝。

因为四膜虫接合对在发育阶段是共享细胞质的，因而研究人员可以观察到这种抑制效应是通过细胞质来传递的，因而只要接合双方的一方在其亲本的体细胞基因组中含有 IES 的同源序列，就可以对后代的移除产生影响。因此，亲本核 DNA 的重排状态是通过细胞质来传递的，并且调控在下代新的体细胞基因组中事件的发生。将接合方一方的状态传递给另一方并不需要生殖细胞基因组的交换，也就排除了一种可能：即传递是通过在发育前正常细胞生长过程中发生生殖细胞的印记完成的（Chalker et al. 2005）。事实上，通过分开接合的双方，一方是野生型细胞而其配偶是体细胞基因组包含多拷贝 M 片段的细胞，发现 DNA 状态传递只在减数分裂以后发生，与发育中的核首次开始分化成新的大核和小核的时间非常接近。因此，母体体细胞基因组的影响在整个发育过程中是很活跃的。

### 母体控制的生物学意义

纤毛生物具有的改变 DNA 重排的能力反应了母体的模式通过一种简单的方式将不同的体细胞 DNA 模式传递给后代的细胞。这种动态的调控表明了两种不同的遗传状态的平衡，例如 $IES^+$ 和 $IES^-$，这种平衡达到并可以在很多代间得到维持。有实验显示至少一个草履虫的 IES 的母体影响足以使得后代中一小部分大核呈 $IES^+$ 的细胞转变成含有较大比例的大核 $IES^-$ 的细胞（Duharcourt et al. 1995）。这种影响也与四膜虫 O 与 E

接合型的稳定传代相关。当然在两个系统中都发现了一定程度的不对称性，因为 IES 移除和 O 型接合型的决定是缺省的发育通路，而 IES 的保持和 E 型的决定需要细胞质的信号来改变既定的通路。这些系统会很好的相关，因为孟德尔 $mtF^E$ 突变体（削弱了 *G* 基因的 IES 的移除）在不考虑亲本状态时会引导 E 接合型的决定，这个发现将两个系统联系到了一起。表明在这些现象背后的机制毫无疑问揭示了表观遗传信息的母系继承的全新模式。

## 11. 由 RNA 干扰介导的整个基因组的跨核比较

上面提到的同源依赖效应表明了母体体细胞和生殖细胞基因组之间在细胞核的分化中会有信息的相互传递，这种交叉会改变 DNA 的重排模式。只有高度同源性的序列受到影响表明这种交叉是由核酸介导的。这种调控机制可以包含母体大核的转录产物，在大核的发育时期向外运输，如果转录产物中包含 IES，它们可以阻止同源序列的移除。然而，也有实验表明假定的母体的转录产物在 IES 移除引起的双链缺口的修复过程中不会参与成为供体模板（Duharcourt et al. 1995）。如果保护性的母体转录产物存在，那么引起基因沉默的 23 nt 的 siRNA 会间接通过加速降解相同序列的保护性转录产物来达到 DNA 移除作用的目的。仅仅是 siRNA 的这个功能是不能解释重排模式的母系继承的。就像在上面所讨论的，似乎母体大核的转录产物和源于减数分裂小核的一组新类型的短小 RNA 的共同作用会最终介导基因组的重排。

### 11.1 DNA 移除与短 RNA 的关系

在四膜虫中，经典的异染色质修饰 H3K9 甲基化，在 IES 移除前标记了与 IES 相关的染色质。这个修饰是如何特异性地标记这些最终将被移除的 DNA 片段的呢？有人认为一条类似 RNAi 的通路通过 RNA 来介导它的正确定位（综述见 Mochizuki and Gorovsky 2004；Yao and Chao 2005）。利用同源 RNA 分子来移除 DNA 的最初预想是通过发现四膜虫 IES 在接合早期是双向转录的现象而得到启发的（Chalker and Yao 2001）。真正的突破是 Mochizuki 等（2002）发现了 PIWI/Argonaute 家族蛋白（由 *TWI1* 编码）在整个发育过程中表达，并对 DNA 的重排作用非常关键（Mochizuki et al. 2002）。因为 Argonaute 蛋白是 RNAi 介导过程的主要参与者，这些研究者寻找并发现了一组内源的短小 RNA（约 28 nt），这些 RNA 能与生殖细胞基因组中序列特异性的互补。这些小 RNA 的干扰会破坏 *TWI1* 基因，使得发育体细胞核中的 H3K9 甲基化现象消失。这项工作发表于 2002 年，与裂殖酵母中的实验（第 6、8 章）共同建立了一个新的理论，即异染色质是由小的同源性的 RNA 以 H3K9 双甲基化沉默为目标标记到特定的位点而形成的。

*DCL1* 基因编码 Dicer 核糖核酸酶会将双向转录产生的 RNA 加工成约 28 nt 的小 RNA，*DCL1* 的发现为整个过程的调控提供了新的视野（Malone et al. 2005；Mochizuki and Gorovsky 2005）。该蛋白在接合早期高表达，定位于即将进行减数分裂的小核中，表明小 RNA 的产生是暂时的，并且空间上与生殖细胞核相关。*DCL1* 的干扰会导

致小 RNA 无法产生，生殖细胞基因组的转录产物的累计，最终导致 IES 移除无法完成。令人惊奇的是，小 RNA 的缺失不会导致影响 H3K9 甲基化，就像在 *TWI1* 敲除的细胞系中所看到的。尽管如此，染色质免疫沉淀分析表明这种修饰不再集中于 IES 区域；因此，小 RNA 是这种染色质修饰分布在正确位点上的保证（Malone et al. 2005）。

为了解释母体基因组在调控这些事件中的作用，扫描 RNA(“scan” RNA，scnRNA)模型被提出（Mochizuki et al. 2002）。图 7-9 中展示的这个模型，28 nt 的扫描 RNA 是在小核中产生的，与包含类 RISC 复合体的 Twi1 蛋白组装在一起分布于细胞质中，这最初是运向大核的。在那里他们“扫描”已经存在的同源的重排基因。与母体序列配对的 scnRNA 从活性复合体中脱落下来，剩余的 Twi1-scnRNA 复合体被转运到发育的大核处并使 H3K9 甲基化定位于同源的序列上，为它们的移除做标记，完成 DNA 的重排。这个模型后来被另一发现证实：即 Twi1 在发育早期，大量小 RNA 产生之后是定位在母体大核上，但是在新的大核前体出现之后便不是了。因此，小 RNA 蛋白复合体的调控转运方便了体细胞和生殖细胞基因组的比较。

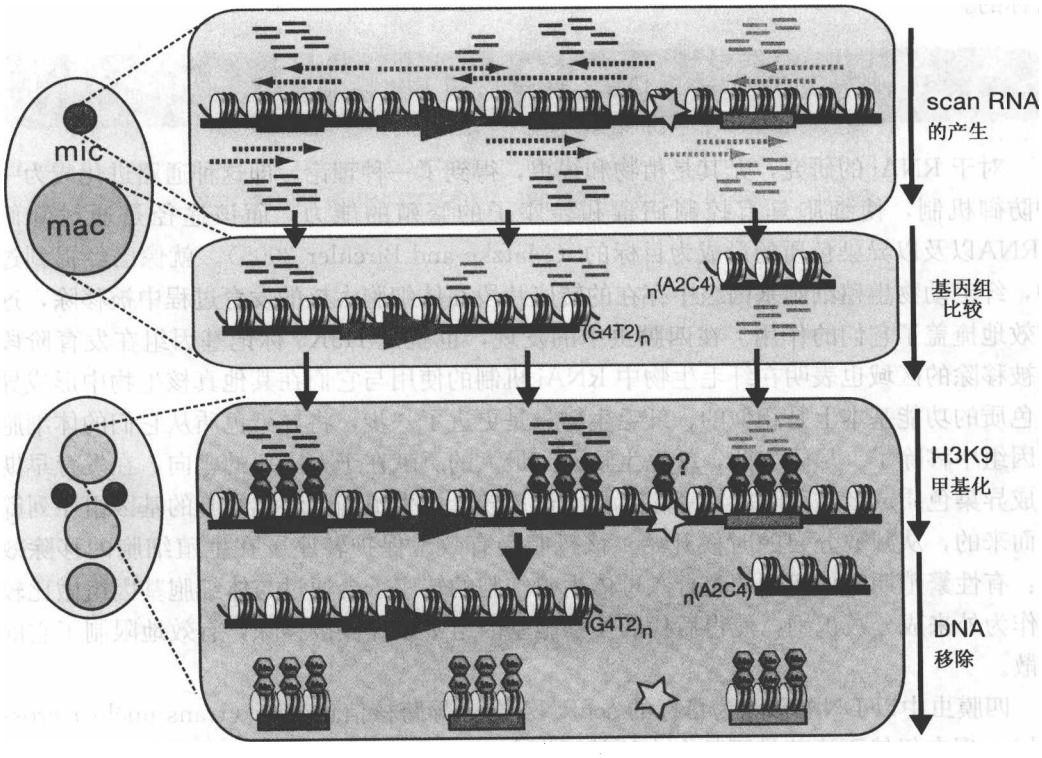



图 7-9　DNA 切除控制的 RNA 扫描模式

大部分生殖细胞基因组的双向转录发生在发育早期并导致 scnRNA 的产生。scnRNA 进而被运到母本大核内。在此若遇同源序列将导致它们从激活池中的移除。余下的 scnRNA 被再次靶向发育中的大核，并以 H3K9 甲基化为目标定位于同源序列，进而通过信号作用从基因组中移除该序列（该模型修改自 Mochizuki et al. 2002）。

## 11.2 在草履虫中将 RNA 从母体转移到合子大核中

生殖细胞和体细胞基因组的序列对比会需要十分复杂的机制，同时保证大量 RNA 分子在核间的转运以及由扫描模型提示的对大数量的配对的影响。新发现的核酸结合蛋白 Nowa1 似乎在草履虫中参与到这种核间的信息传递（Nowacki et al. 2005）。Nowa1 在减数分裂以前合成，首先在母体大核中累计，就像 Twi1 蛋白一样，然后在合子大核发育时期重新定位。在 Nowa1 蛋白后连接一个光激发的 GFP 使得研究人员可以下结论说这种蛋白是从一种核型转运到另一种上的。该蛋白的一个结构域可以结合 RNA，而第二个结构域是核间转运必须、必要的；因此，Nowa1 可能是一个 RNA 转运者。

Nowa1 在新的大核发育中十分重要，包括生殖细胞转座子的移除和 IES 子集的移除。令人惊奇的是，只有那些受母体调控的 IES 是受到 Nowa1 敲除影响的，表明蛋白本身是参与核间基因组比对的。对同源序列存在不敏感的 IES 并不依赖于 Nowa1 完成它的移除，更确认了草履虫体内存在机制不同的 IES。尽管体内与 Nowa1 结合的核酸还不知晓，Nowa1 缺失时的影响与提出的扫描理论是吻合的（图 7-9），也表明 Nowa1 蛋白可能携带 RNA，而这些 RNA 是在发育的大核中以受母体控制的 IES 和转座子为目标的。

## 12. 结论：DNA 移除是基因组防御机制

对于 RNAi 的研究，尤其是植物和线虫，得到了一种理论，即这种通路进化成为一种防御机制，使细胞具有控制病毒和转座子的繁殖的能力，而该途径是通过降解 mRNA以及以异染色质的形成为目标的（Matzke and Birchler 2005）。就像已经提到过的，纤毛动物生殖细胞基因组中存在的转座片段在体细胞大核的发育过程中被移除，这有效地掩盖了它们的作用。在四膜虫中的发现，也就是 H3K9 标记基因组在发育阶段会被移除的区域也表明在纤毛生物中 RNAi 机制的使用与它们在其他真核生物中形成异染色质的功能基本上是相似的；纤毛生物只是更进了一步，将异染色质从它们的体细胞基因组中移除了。尽管如此，纤毛生物研究最大的贡献在于 RNAi 的靶向，在发育早期形成异染色质或者被移除的特意序列是通过比较母体生殖细胞和体细胞的基因组序列筛选而来的，从减数分裂的时候开始。该机制会有效地保护转座子在生殖细胞的移除效应：有性繁殖期间任何新的整合入母体生殖细胞的转座子会通过与体细胞基因组做比较而作为外来成分被识别，使得后代的体细胞基因组中该片段被移除，有效地限制了它的扩散。

四膜虫中以 DNA 移除为目标的 scnRNA 会介导跨核信息交流（trans-nuclear cross talk）。现在仍然不清楚是否整个小核基因组产生 scnRNA，但是未发表的结果发现了草履虫中的证据表明至少一大部分是这样的。Northern 印记分析也发现了减数分裂特异的、内源的短小 RNA，这些 RNA 与细胞的基因相对应，也与转座子和 IES 对应（M. Nowacki et al. 未发表）。这些短小 RNA 长 25nt，与 siRNA 的 23 nt 明显的区分开来。他们的功能上与四膜虫的 scnRNA 相似，因为草履虫特异的类 *Dicer* 基因的失活会导致

它们在减数分裂时期无法产生，无法移除转座子和母体控制的 IES，同样也无法将母系继承的大核基因的缺失传递到下一代（V. Serrano et al. 未发表）。

最初的基因组扫描模型假设 scnRNA 是与大核中重排的 DNA 相比较后，将那些匹配的大核基因序列失活，并选做小核特异基因池的（图 7-10）。然而，几条证据表明 scnRNA可能与大核衍生的转录产物比较而非 DNA 本身。如果与 DNA 比较则需要将整个基因组 DNA 解旋以进行比较。引入四膜虫母体大核、并可以干扰 DNA 移除的 IES

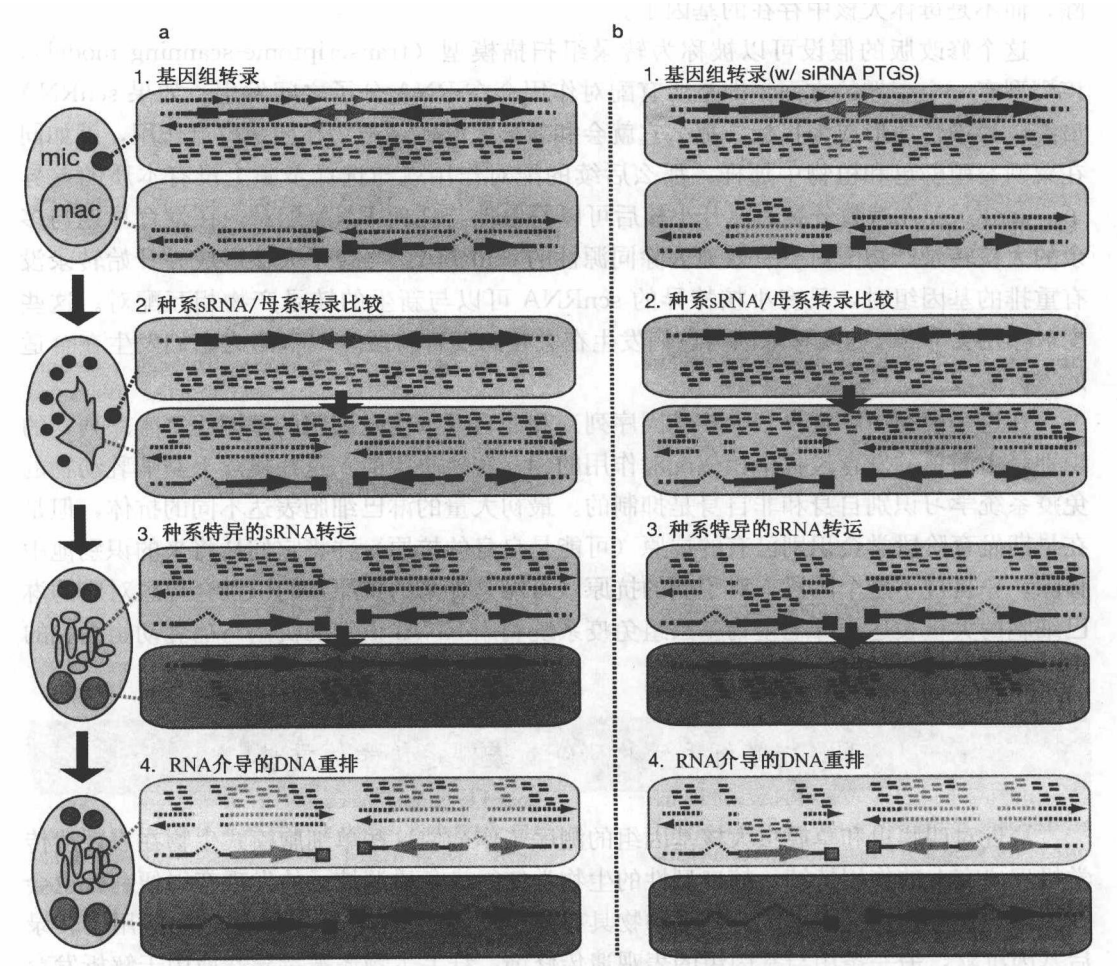



图 7-10　生殖细胞和体细胞基因组比较的转录组扫描模型

(a) 调控缺陷（1）减数分裂开始时，大多数或者几乎所有的大核和小核基因组均发生转录；大核内的虚线表明无特征性转录。小核转录是双向的，产生适用于各种序列的 scnRNA（短双链分子）（细胞基因，亮紫色箭头；转座子，橘黄色双箭头；IES，绿色框）（2）scnRNA 进入母本大核，进而与同源的转录因子配对。配对也可以发生在胞质内。（3）在小核特异的 scnRNA 进入发育中的合子大核时，同源转录因子配对的 scnRNA 有序或无序地和同源序列（DNA 或新生转录子）配对，从而定位 H3K9 甲基化到小核特异序列（转座子和 IES）。（4）标记的序列被移除。(b) 母本大核基因的转录后沉默。通过高拷贝转基因或 dsRNA 的转录后沉默实验诱导产生双链 siRNA 至基因（红色）。这些 siRNA 降解同源母本体细胞转录使得同源 scnRNA 不被失活且游离于合子大核发育中的基因消除的靶标之外。

是经转录的，使得这个假说可能成立（Chalker and Yao 2001；Chalker et al. 2005）。从 IES 衍生的转录产物可能通过降解同源 scnRNA 的途径来保护合子 IES 免受移除。细胞质中存在这样一个因子可以补救 d48 株中 *A* 基因移除的证据也同样支持保护性转录产物的存在，更进一步表明 scnRNA 的选择会发生在胞质中也会发生在母体大核中。最终，这可以解释转录后基因的沉默可以导致它在下一代中的移除：如果母体的大核转录产物被 23 nt 的 siRNA 降解，同源的 scnRNA 就会将目标集中于新大核中的基因移除，而不是母体大核中存在的基因了。

这个修改版的假设可以被称为转录组扫描模型（transcriptome-scanning model），该模型有一个有趣的方面，即是所有配对作用会在 RNA 分子之间发生。如果 scnRNA 最终以异染色质形成为目标，那么它就会和新形成的转录产物在同源位点作用，就如同在粟酒裂殖酵母和植物中那样，那么后续的配对作用就与选择步骤上没有本质的差异了。scnRNA 在减数分裂后离开小核后可以简单地与任何 RNA 配对。在发育早期与多余的大核转录产物配对会有效地去除同源性的 scnRNA。当合子大核形成并开始转录没有重排的基因组时，只有小核特异的 scnRNA 可以与新生的转录产物相互配对。这些配对作用会导致 DNA 移除因为它们发生在大核的发育阶段，或者因为它们发生在合适的发育阶段。

在这个模型中，自身（大核基因序列）和非自身（生殖细胞的基因组寄生序列）的识别是由简单的发育过程中相似配对作用的相互转换决定的。这在概念上与哺乳动物的免疫系统学习识别自身和非自身是抑制的。最初大量的淋巴细胞表达不同的抗体，但是在早期发育阶段学会识别已有的抗原（可能是自身的抗原）并将它们从将来的识别池中移除。一旦过了这个阶段，对于新的抗原（可能是外来的抗原）的识别会导致对应的淋巴细胞的大量表达。纤毛生物基因组免疫系统利用 RNAi 识别的机制与哺乳动物细胞的免疫系统是平行的。

## 13. 未来纤毛动物研究对表观遗传学的贡献

最近对四膜虫和草履虫大核基因组的测序工作会使这些单细胞模式生物在表观遗传学机理研究上的作用持续。核二型性的生物学使得染色质调控子的发现变得便利，也会为更新的发现提供优势系统。纤毛动物具有两条不同的 RNA 干扰通路，一条用于转录后基因沉默，另一条用与基因组的表观遗传修饰，纤毛生物还被特异性地用于解析发育过程中 RNA 指导的，同源依赖的调控过程的复杂性。Dicer 相关的核糖核酸酶，Argonaute 类似的蛋白，还有它们已经发现或者即将被发现的相互作用蛋白，现代研究都致力于研究它们功能的特异性。得出进化树上纤毛生物的系统发育位置。这些信息会有助于了解这些机制在真核生物中的进化过程。

正是不同表观遗传现象的研究使得人们承认了真核生物中 RNA 介导的调控。但是，甚至在这层了解之前，一个核的 DNA 会影响到另一核同源序列的发现就促使纤毛生物学家提出了反式作用的 RNA 分子会指导同源依赖型核酸信息交流的假设。RNA 干扰和它在 DNA 重排中的作用证实了这种假设。未来对这种普遍联系的深入了解可能

就源自于解析是什么介导了跨核信息交流的。

如果“转录组扫描”假设（图 7-10）是正确的，那么母体体细胞的转录产物阻止了同源合子 DNA 序列的移除。如果将 DNA 移除与异染色质的形成等同看，那么母体转录产物通过阻抑 DNA 的移除来促进常染色质同源序列特异性的修饰。那么这些转录产物对同源基因的作用是不直接的，它们的作用途径是将同源的 scnRNA 失活，因为 scnRNA 会协助异染色质的形成。然而，母体转录产物的这种作用代表了一种新的机理即所有已知的 RNA 介导的、同源依赖的机理会引起下游目的基因的调控，而这些母体转录产物不一定是编码蛋白的 mRNA。本质上说，小 RNA 被较长的转录产物降解是与已经证实的基于 RNA 干扰的机理相违背的，又凭什么说是小 RNA 使长转录产物失活呢？

如果长和短 RNA 的作用相互拮抗，那么它们的相互作用会使得系统朝两个方向中的一个推进，决定于它们相对的富余量。这表明基因表达状态的遗传可以依赖于来自 RNA 群体的持续反馈，而不是仅仅依赖于半保留的染色质复制机理。这很容易让人相信这种 RNA 介导的、反式作用的机制是决定了草履虫无性生殖过程中表面抗原基因表达的稳定性的，因为这同样可以解释在完成有性生殖以后，发育中的大核是如何将母体的血清型通过细胞质传给后代的。因而，基因组重排模式的表观遗传可能只是复杂的基于 RNA 的遗传系统的一个部分而已。

在将来的纤毛动物研究进一步解释 RNA 和蛋白如何介导调控的同时，这种机制在其他生物中的存在以及它与已知的表观遗传机制的联系会变得越来越明朗。基于 RNA 的遗传机制是广泛存在的这个设想并非不可信的。近期对于多核真核生物转录组向外运输的研究表明，其体内存在大量且复杂的非编码转录产物（Mattick 2004；Meyers et al. 2004；Suzuki and Hayashizaki 2004；Cheng et al. 2005；Clacerie 2005），同时也发现了一些与所有种类的基因组都匹配的小 RNA 的普遍存在（Lu et al. 2005）。很多同源依赖性的作用，包括真菌的减数分裂期基因沉默和植物中的副突变现象都是没有得到完全解析的（见第 6、9 章）。这需要所有真核生物当然也包括纤毛动物的实验结果的融合、交叉，或许这样才能全方位地揭示表观遗传学。

赵　璟　王大亮　译
王大亮　校
孙方霖　审校

## 参考文献

Allis C.D. and Gorovsky M.A. 1981. Histone phosphorylation in macro- and micronuclei of *Tetrahymena thermophila*. *Biochemistry* **20:** 3828–3833.

Allis C.D., Glover C.V., Bowen J.K., and Gorovsky M.A. 1980. Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, *Tetrahymena thermophila*. *Cell* **20:** 609–617.

Beisson J. and Sonneborn T.M. 1965. Cytoplasmic inheritance of the organization of the cell cortex in *Paramecium aurelia*. *Proc. Natl. Acad. Sci.* **53:** 275–282.

Beisson J. and Wright M. 2003. Basal body/centriole assembly and continuity. *Curr. Opin. Cell Biol.* **15:** 96–104.

Betermier M. 2004. Large-scale genome remodelling by the developmentally programmed elimination of germ line sequences in the ciliate *Paramecium*. *Res. Microbiol.* **155:** 399–408.

Betermier M., Duharcourt S., Seitz H., and Meyer E. 2000. Timing of developmentally programmed excision and circularization of *Paramecium* internal eliminated sequences. *Mol. Cell. Biol.* **20:** 1553–1561.

Brownell J.E. and Allis C.D. 1995. An activity gel assay detects a sin-

gle, catalytically active histone acetyltransferase subunit in *Tetrahymena* macronuclei. *Proc. Natl. Acad. Sci.A* **92:** 6364–6368.

Brownell J.E., Zhou J., Ranalli T., Kobayashi R., Edmondson D.G., Roth S.Y., and Allis C.D. 1996. *Tetrahymena* histone acetyltransferase: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84:** 843–851.

Brygoo Y. and Keller A.-M. 1981a. A mutation with pleitropic effects on macronuclearly differentiated functions in *Paramecium tetraurelia*. *Dev. Genet.* **2:** 23–34.

———. 1981b. Genetic analysis of mating type differentiation in *Paramecium tetraurelia*. III. A mutation restricted to mating type *E* and affecting the determination of mating type. *Dev. Genet.* **2:** 13–22.

Cassidy-Hanley D., Bisharyan Y., Fridman V., Gerber J., Lin C., Orias E., Orias J.D., Ryder H., Vong L., and Hamilton E.P. 2005. Genome-wide characterization of *Tetrahymena thermophila* chromosome breakage sites. II. Physical and genetic mapping. *Genetics* **170:** 1623–1631.

Chalker D.L. and Yao M.C. 1996. Non-Mendelian, heritable blocks to DNA rearrangement are induced by loading the somatic nucleus of *Tetrahymena thermophila* with germ line-limited DNA. *Mol. Cell. Biol.* **16:** 3658–3667.

———. 2001. Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in *Tetrahymena thermophila*. *Genes Dev.* **15:** 1287–1298.

Chalker D.L., Fuller P., and Yao M.C. 2005. Communication between parental and developing genomes during *Tetrahymena* nuclear differentiation is likely mediated by homologous RNAs. *Genetics* **169:** 149–160.

Cheng J., Kapranov P., Drenkow J., Dike S., Brubaker S., Patel S., Long J., Stern D., Tammana H., Helt G., et al. 2005. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* **308:** 1149–1154.

Claverie J.M. 2005. Fewer genes, more noncoding RNA. *Science* **309:** 1529–1530.

Cleveland D.W., Mao, Y., and Sullivan K.F. 2003. Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. *Cell* **112:** 407–421.

Coyne R.S., Chalker D.L., and Yao M.C. 1996. Genome downsizing during ciliate development: Nuclear division of labor through chromosome restructuring. *Annu. Rev. Genet.* **30:** 557–578.

Coyne R.S., Nikiforov M.A., Smothers J.F., Allis C.D., and Yao M.C. 1999. Parental expression of the chromodomain protein Pdd1p is required for completion of programmed DNA elimination and nuclear differentiation. *Mol. Cell* **4:** 865–872.

Davis M.C., Ward J.G., Herrick G., and Allis C.D. 1992. Programmed nuclear death: Apoptotic-like degradation of specific nuclei in conjugating *Tetrahymena*. *Dev. Biol.* **154:** 419–432.

Dou Y. and Gorovsky M.A. 2000. Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch. *Mol. Cell* **6:** 225–231.

———. 2002. Regulation of transcription by H1 phosphorylation in *Tetrahymena* is position independent and requires clustered sites. *Proc. Natl. Acad. Sci.* **99:** 6142–6146.

Dou Y., Mizzen C.A., Abrams M., Allis C.D., and Gorovsky M.A. 1999. Phosphorylation of linker histone H1 regulates gene expression in vivo by mimicking H1 removal. *Mol. Cell* **4:** 641–647.

Duharcourt S., Butler A., and Meyer E. 1995. Epigenetic self-regulation of developmental excision of an internal eliminated sequence on *Paramecium tetraurelia*. *Genes Dev.* **9:** 2065–2077.

Duharcourt S., Keller A.M., and Meyer E. 1998. Homology-dependent maternal inhibition of developmental excision of internal eliminated sequences in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **18:** 7075–7085.

Epstein L.M. and Forney J.D. 1984. Mendelian and non-Mendelian mutations affecting surface antigen expression in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **4:** 1583–1590.

Fan Q. and Yao M.-C. 1996. New telomere formation coupled with site-specific chromosome breakage in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **16:** 1267–1274.

Forney J.D. and Blackburn E.H. 1988. Developmentally controlled telomere addition in wild-type and mutant paramecia. *Mol. Cell. Biol.* **8:** 251–258.

Forney J.D., Yantiri F., and Mikami K. 1996. Developmentally controlled rearrangement of surface protein genes in *Paramecium tetraurelia*. *J. Eukaryot. Microbiol.* **43:** 462–467.

Galvani A. and Sperling L. 2001. Transgene-mediated post-transcriptional gene silencing is inhibited by 3' non-coding sequences in *Paramecium*. *Nucleic Acids Res.* **29:** 4387–4394.

———. 2002. RNA interference by feeding in *Paramecium*. *Trends Genet.* **18:** 11–12.

Garnier O., Serrano V., Duharcourt S., and Meyer E. 2004. RNA-mediated programming of developmental genome rearrangements in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **24:** 7370–7379.

Gratias A. and Betermier M. 2001. Developmentally programmed excision of internal DNA sequences in *Paramecium aurelia*. *Biochimie* **83:** 1009–1022.

———. 2003. Processing of double-strand breaks is involved in the precise excision of *Paramecium* internal eliminated sequences. *Mol. Cell. Biol.* **23:** 7152–7162.

Hamilton E., Bruns P., Lin C., Merriam V., Orias E., Vong L., and Cassidy-Hanley D. 2005. Genome-wide characterization of *Tetrahymena thermophila* chromosome breakage sites. I. Cloning and identification of functional sites. *Genetics* **170:** 1611–1621.

Harumoto T. 1986. Induced change in a non-Mendelian determinant by transplantation of macronucleoplasm in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **6:** 3498–3501.

Harwood J. 1985. The erratic career of cytoplasmic inheritance. *Trends Genet.* **1:** 298–300.

Hayashi T., Hayashi H., Fusauchi Y., and Iwai K. 1984. *Tetrahymena* histone H3. Purification and two variant sequences. *J. Biochem.* **95:** 1741–1749.

Hayashi T., Hayashi H., and Iwai K. 1987. *Tetrahymena* histone H1. Isolation and amino acid sequence lacking the central hydrophobic domain conserved in other H1 histones. *J. Biochem.* **102:** 369–376.

Jahn C.L. and Klobutcher L.A. 2002. Genome remodeling in ciliated protozoa. *Annu. Rev. Microbiol.* **56:** 489–520.

Jaraczewski J.W. and Jahn C.L. 1993. Elimination of Tec elements involves a novel excision process. *Genes Dev.* **7:** 95–105.

Jessop-Murray H., Martin L.D., Gilley D., Preer J.R., and Polisky B. 1991. Permanent rescue of a non-Mendelian mutation of *Paramecium* by microinjection of specific DNA sequences. *Genetics* **129:** 727–734.

Klobutcher L.A. 1999. Characterization of in vivo developmental chromosome fragmentation intermediates in *E. crassus*. *Mol. Cell* **4:** 695–704.

Klobutcher L.A. and Herrick G. 1995. Consensus inverted terminal repeat sequence of *Paramecium* IESs: Resemblance to termini of Tc1-related and *Euplotes* Tec transposons. *Nucleic Acids Res.* **23:** 2006–2013.

———. 1997. Developmental genome reorganization in ciliated protozoa: The transposon link. *Prog. Nucleic Acid Res. Mol. Biol.* **56:** 1–62.

Klobutcher L.A., Turner L.R., and LaPlante J. 1993. Circular forms of developmentally excised DNA in *Euplotes crassus* have a heteroduplex junction. *Genes Dev.* **7:** 84–94.

Kobayashi S. and Koizumi S. 1990. Characterization of non-Mendelian and Mendelian mutant strains by micronuclear transplantation in *Paramecium tetraurelia*. *J. Protozool.* **37:** 489–492.

Koizumi S. and Kobayashi S. 1989. Microinjection of plasmid DNA encoding the A surface antigen of *Paramecium tetraurelia* restores the ability to regenerate a wild-type macronucleus. *Mol. Cell. Biol.* **9:** 4398–4401.

Landweber L.F., Kuo T.C., and Curtis E.A. 2000. Evolution and assembly of an extremely scrambled gene. *Proc. Natl. Acad. Sci.* **97:** 3298–3303.

Le Mouël A., Butler A., Caron F., and Meyer E. 2003. Developmentally regulated chromosome fragmentation linked to imprecise elimination of repeated sequences in Paramecia. *Eukaryot. Cell* **2:** 1076–1090.

Lee S.R. and Collins K. 2006. Two classes of endogenous small RNAs in *Tetrahymena thermophila*. *Genes Dev.* **20:** 28–33.

Liu Y., Mochizuki K., and Gorovsky M.A. 2004. Histone H3 lysine 9

methylation is required for DNA elimination in developing macronuclei in *Tetrahymena*. *Proc. Natl. Acad. Sci.* **101:** 1679–1684.

Liu Y., Song X., Gorovsky M.A., and Karrer K.M. 2005. Elimination of foreign DNA during somatic differentiation in *Tetrahymena thermophila* shows position effect and is dosage dependent. *Eukaryot. Cell* **4:** 421–431.

Lu C., Tej S.S., Luo S., Haudenschild C.D., Meyers B.C., and Green P.J. 2005. Elucidation of the small RNA component of the transcriptome. *Science* **309:** 1567–1569.

Madireddi M.T., Davis M., and Allis D. 1994. Identification of a novel polypeptide involved in the formation of DNA-containing vesicles during macronuclear development in *Tetrahymena*. *Dev. Biol.* **165:** 418–431.

Madireddi M.T., Coyne R.S., Smothers J.F., Mickey K.M., Yao M.C., and Allis C.D. 1996. Pdd1p, a novel chromodomain-containing protein, links heterochromatin assembly and DNA elimination in *Tetrahymena*. *Cell* **87:** 75–84.

Malone C.D., Anderson A.M., Motl J.A., Rexer C.H., and Chalker D.L. 2005. Germ line transcripts are processed by a Dicer-like protein that is essential for developmentally programmed genome rearrangements of *Tetrahymena thermophila*. *Mol. Cell. Biol.* **25:** 9151–9164.

Mattick J.S. 2004. RNA regulation: A new genetics? *Nat. Rev. Genet.* **5:** 316–323.

Matzke M.A. and Birchler J.A. 2005. RNAi-mediated pathways in the nucleus. *Nat. Rev. Genet.* **6:** 24–35.

Meyer E. 1992. Induction of specific macronuclear developmental mutations by microinjection of a cloned telomeric gene in *Paramecium primaurelia*. *Genes Dev.* **6:** 211–222.

Meyer E. and Keller A.M. 1996. A Mendelian mutation affecting mating-type determination also affects developmental genomic rearrangements in *Paramecium tetraurelia*. *Genetics* **143:** 191–202.

Meyer E., Butler A., Dubrana K., Duharcourt S., and Caron F. 1997. Sequence-specific epigenetic effects of the maternal somatic genome on developmental rearrangements of the zygotic genome in *Paramecium primaurelia*. *Mol. Cell. Biol.* **17:** 3589–3599.

Meyers B.C., Vu T.H., Tej S.S., Ghazal H., Matvienko M., Agrawal V., Ning J., and Haudenschild C.D. 2004. Analysis of the transcriptional complexity of *Arabidopsis thaliana* by massively parallel signature sequencing. *Nat. Biotechnol.* **22:** 1006–1011.

Mizzen C.A., Dou Y., Liu Y., Cook R.G., Gorovsky M.A., and Allis C.D. 1999. Identification and mutation of phosphorylation sites in a linker histone. Phosphorylation of macronuclear H1 is not essential for viability in *Tetrahymena*. *J. Biol. Chem.* **274:** 14533–14536.

Mochizuki K. and Gorovsky M.A. 2004. Small RNAs in genome rearrangement in *Tetrahymena*. *Curr. Opin. Genet. Dev.* **14:** 181–187.

———. 2005. A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase. *Genes Dev.* **19:** 77–89.

Mochizuki K., Fine N.A., Fujisawa T., and Gorovsky M.A. 2002. Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. *Cell* **110:** 689–699.

Moreira-Leite F.F., Sherwin T., Kohl L., and Gull K. 2001. A trypanosome structure involved in transmitting cytoplasmic information during cell division. *Science* **294:** 610–612.

Nikiforov M.A., Gorovsky M.A., and Allis C.D. 2000. A novel chromodomain protein, pdd3p, associates with internal eliminated sequences during macronuclear development in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **20:** 4128–4134.

Nowacki M., Zagorski-Ostoja W., and Meyer E. 2005. Nowa1p and Nowa2p: Novel putative RNA binding proteins involved in trans-nuclear crosstalk in *Paramecium tetraurelia*. *Curr. Biol.* **15:** 1616–1628.

Paschka A.G., Jönsson F., Maier V., Möllenbeck M., Paeschke K., Postberg J., Rupprecht S., and Lipps H.J. 2003. The use of RNAi to analyze gene function in spirotrichous ciliates. *Eur. J. Protistol.* **39:** 449–454.

Philippe H., Germot A., and Moreira D. 2000. The new phylogeny of eukaryotes. *Curr. Opin. Genet. Dev.* **10:** 596–601.

Pond F.R., Gibson I., Lalucat J., and Quackenbush R.L. 1989. R-body-producing bacteria. *Microbiol. Rev.* **53:** 25–67.

Preer J.R., Jr., Preer L.B., and Jurand A. 1974. Kappa and other endosymbionts in *Paramecium aurelia*. *Bacteriol. Rev.* **38:** 113–163.

Prescott D.M. 1994. The DNA of ciliated protozoa. *Microbiol. Rev.* **58:** 233–267.

———. 1999. The evolutionary scrambling and developmental unscrambling of germline genes in hypotrichous ciliates. *Nucleic Acids Res.* **27:** 1243–1250.

Prescott D.M., Ehrenfeucht A., and Rozenberg G. 2003. Template-guided recombination for IES elimination and unscrambling of genes in stichotrichous ciliates. *J. Theor. Biol.* **222:** 323–330.

Ruiz F., Vayssie L., Klotz C., Sperling L., and Madeddu L. 1998. Homology-dependent gene silencing in *Paramecium*. *Mol. Biol. Cell* **9:** 931–943.

Saveliev S.V. and Cox M.M. 1995. Transient DNA breaks associated with programmed genomic deletion events in conjugating cells of *Tetrahymena thermophila*. *Genes Dev.* **9:** 248–255.

———. 1996. Developmentally programmed DNA deletion in *Tetrahymena thermophila* by a transposition-like reaction pathway. *EMBO J.* **15:** 2858–2869.

Scott J.M., Mikami K., Leeck C.L., and Forney J.D. 1994. Non-Mendelian inheritance of macronuclear mutations is gene specific in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **14:** 2479–2484.

Shen X. and Gorovsky M.A. 1996. Linker histone H1 regulates specific gene expression but not global transcription in vivo. *Cell* **86:** 475–483.

Shen X., Yu L., Weir J.W., and Gorovsky M.A. 1995. Linker histones are not essential and affect chromatin condensation in vivo. *Cell* **82:** 47–56.

Shorter J. and Lindquist S. 2005. Prions as adaptive conduits of memory and inheritance. *Nat. Rev. Genet.* **6:** 435–450.

Sonneborn T.M. 1937. Sex, sex inheritance and sex determination in *Paramecium aurelia*. *Proc. Natl. Acad. Sci.* **23:** 378–385.

———. 1975. *Paramecium aurelia*. In *Handbook of genetics* (ed. R. King), pp. 469–594. Plenum, New York.

———. 1977. Genetics of cellular differentiation: Stable nuclear differentiation in eucaryotic unicells. *Annu. Rev. Genet.* **11:** 349–367.

Strahl B.D., Ohba R., Cook R.G., and Allis C.D. 1999. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in *Tetrahymena*. *Proc. Natl. Acad. Sci.* **96:** 14967–14972.

Suzuki M. and Hayashizaki Y. 2004. Mouse-centric comparative transcriptomics of protein coding and non-coding RNAs. *Bioessays* **26:** 833–843.

Taverna S.D., Coyne R.S., and Allis C.D. 2002. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in *Tetrahymena*. *Cell* **110:** 701–711.

Timmons L. and Fire A. 1998. Specific interference by ingested dsRNA. *Nature* **395:** 854.

Timmons L., Court D.L., and Fire A. 2001. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263:** 103–112.

Vavra K.J., Allis C.D., and Gorovsky M.A. 1982. Regulation of histone acetylation in *Tetrahymena* macro- and micronuclei. *J. Biol. Chem.* **257:** 2591–2598.

Wei Y., Yu L., Bowen J., Gorovsky M.A., and Allis C.D. 1999. Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell* **97:** 99–109.

Wei Y., Mizzen C.A., Cook R.G., Gorovsky M.A., and Allis C.D. 1998. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in *Tetrahymena*. *Proc. Natl. Acad. Sci.* **95:** 7480–7484.

Williams K., Doak T.G., and Herrick G. 1993. Developmental precise excision of *Oxytricha trifallax* telomere-bearing elements and formation of circles closed by a copy of the flanking target duplication. *EMBO J.* **12:** 4593–4601.

Wu M., Allis C.D., Richman R., Cook R.G., and Gorovsky M.A. 1986. An intervening sequence in an unusual histone H1 gene of *Tetrahymena thermophila*. *Proc. Natl. Acad. Sci.* **83:** 8674–8678.

Yao M.C. and Chao J.L. 2005. RNA-guided DNA deletion in *Tetrahymena:* An RNAi-based mechanism for programmed genome rearrangements. *Annu. Rev. Genet.* **39:** 537–559.

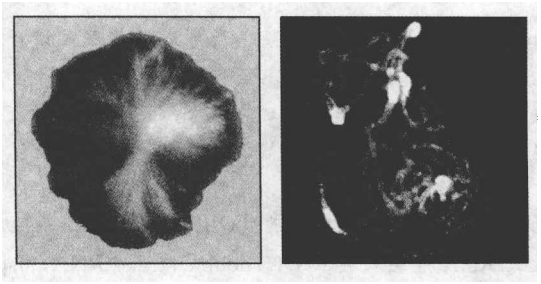
Yao M.C., Duharcourt S., and Chalker D.L. 2002. Genome-wide rearrangements of DNA in ciliates. in *Mobile DNA II* (ed. A. Lambowitz), pp. 730–758. Academic Press, New York.

Yao M.C., Fuller P., and Xi X. 2003. Programmed DNA deletion as an RNA-guided system of genome defense. *Science* **300:** 1581–1584.

You Y., Aufderheide K., Morand J., Rodkey K., and Forney J. 1991. Macronuclear transformation with specific DNA fragments controls the content of the new macronuclear genome in *Paramecium tetraurelia*. *Mol. Cell. Biol.* **11:** 1133–1137.

Yu L. and Gorovsky M.A. 1997. Constitutive expression, not a particular primary sequence, is the important feature of the H3 replacement variant hv2 in *Tetrahymena thermophila*. *Mol. Cell. Biol.* **17:** 6303–6310.

# 第 8 章



# RNAi 和异染色质组装

Robert Martienssen[1] and Danesh Moazed[2]

[1]*Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724*
[2]*Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115-5730*

## 本章目录

1. RNAi 通路的概述
2. 关于 RNA 介导基因转录水平沉默的早期证据
3. 酵母中的 RNAi 和异染色质组装
4. 小 RNA 与 RNAi 效应蛋白复合体共同起始异染色质组装
5. dsRNA 的合成和 siRNA 的产生
6. RNA-RNA 和 RNA-DNA 识别模型
7. RNAi 如何募集染色质修饰酶?
8. 拟南芥中 RNAi 介导的染色质和 DNA 修饰
9. RNAi 介导的染色质修饰在动物中的保守性
10. 小结

参考文献

## 概要

RNA 干扰（RNA interference，RNAi）和异染色质形成的相互交错将这两个基因调控领域联系在了一起，而它们之前被认为是通过不同甚至完全不相关的机制来控制的。早在 80 年前，人们就通过细胞染色的方法发现了异染色质，这些区域在整个细胞周期中一直保持浓缩状态。早期的研究人员在研究染色体结构和基因表达关系的时候就注意到，特定的染色体重排能够导致异染色质结构蔓延至邻近基因，接着这些基因就沉默了。这种蔓延的随机性造成了具有完全相同遗传学背景的细胞群有了不同的表型，这正为我们研究表观遗传调控提供了一个极好的例子。“RNAi”一词最先被用于描述同源反义 RNA 链或双链 RNA（double-strand RNA，dsRNA）导入线虫后引起基因沉默的现象。之前一个解释牵牛花和其他植物中发现的转基因转录后沉默（posttranscriptional transgene silencing，PTGS）现象的类似机制也随即得到了承认。而异染色质则被认为可以广泛地在染色质水平通过基因转录水平沉默（transcriptional gene silencing，TGS）机制直接起作用，导致转录抑制。本章将着重讨论 RNAi 通路和在特定染色体区域形成表观遗传学上可继承的异染色质之间的关系，以及最近在酿酒酵母和拟南芥中发现的能证明这种联系的几个例子。

酿酒酵母的基因组由 3.5Mb～5.7Mb 的三条染色体组成。每条染色体都包含大量的重复 DNA，特别是在包装成异染色质的着丝粒区。另外，接合型位点（此位点决定了细胞类型）和亚端粒 DNA 区域也包含有包装成异染色质的重复序列。现在我们知道 DNA 组装成异染色质有调节性和结构性的作用。在酵母接合型位点中，异染色质调控基因转录对细胞类型是很重要的。在端粒和着丝粒中，异染色质起着结构性作用，这对细胞分裂中染色体的正确分离至关重要。此外，在许多真核细胞基因组中，重复 DNA 序列和转座元件占了很大一部分，在有些例子中甚至占到了一半以上。异染色质及其相关机制在维持基因组稳定中起到关键作用，而这种作用是通过调控重复序列活性来实现的。最近的研究揭示，酿酒酵母中异染色质形成过程竟然需要 RNAi 通路组分的参与，这让我们对“这两种机制如何能够在染色质水平协同作用”这个问题有了一个深入的认识。简单地说，小干扰 RNA（small interfering RNA，siRNA）——RNAi 和其他 dsRNA 沉默机制中的标志物，组装到 RNA 诱导转录沉默复合体（RNA-induced transcriptional silencing，RITS）中，在与其互补的染色体区域指导染色质表观遗传修饰和异染色质形成。RITS 通过 siRNA 的碱基配对来指导其与沉默目标区域的 DNA 或新生 RNA 序列结合，RITS 直接结合甲基化组蛋白 H3 也稳定了这种结合。RITS 的这两个活性与其他著名的异染色质结合因子一起引发了异染色质的形成，从而将 RNA 沉默和异染色质形成直接联系在了一起。

在拟南芥和其他真核生物中（酿酒酵母除外），着丝粒区 DNA 也是由重复元件组成的。这些重复元件和其他重复序列，如逆转录因子和转座子，都是 siRNA 的来源，而 siRNA 又导致 H3K9 和 DNA 的甲基化。启动和维持这些抑制性的甲基化需要 RNAi 通路的几个组分参与。在本章中，我们将讨论酿酒酵母和拟南芥中异染色质的 siRNA 是如何产生并介导 DNA 和（或）染色质修饰的。

## 1. RNAi 通路的概述

尽管“RNAi”一词最先是被用来描述线虫中外源 dsRNA 诱导的基因沉默现象(Fire et al. 1998)，但是现在它更广泛地表示各种 dsRNA 引发的基因沉默。RNAi 涉及的步骤包括：dsRNA 的产生（可以是内源的也可以是外源的，比如病毒 RNA 就是外源的)，加工成 siRNA，定位这些小分子到 mRNA（PTGS）或染色质区域（TGS)，从而引起沉默。因此，我们在介绍 TGS 特异的 RNAi 装置的组分之前，我们先讨论一下启动 RNAi 装置的 dsRNA 的来源。

dsRNA 可以由重复 DNA 元件双向转录产生，也可由能够形成内部碱基配对的 RNA 转录物形成（分别见图 8-1a 和 b)。举个例子来说，反向重复序列区转录产生的 RNA 分子就能够自身折叠形成发卡结构。dsRNA 反向被 Dicer 剪切产生 siRNA，Dicer 是 RNase III 型的核酸酶。这些都是互补的短双链结构，21～27 个核苷酸长，在双链的 3′端有典型的 2nt 的突出碱基（Hamilton and Baulcombe 1999；Zamore et al. 2000；Bernstein et al. 2001；Elbashir et al. 2001；Hannon 2002；Zamore 2002；Bartel 2004；Baulcombe 2004)。这些双链被解开成单链 siRNA，与互补目标序列通过碱基配对起到靶向作用。因此，它们是在 RNAi 介导的沉默机制里起中心作用的特殊因子。

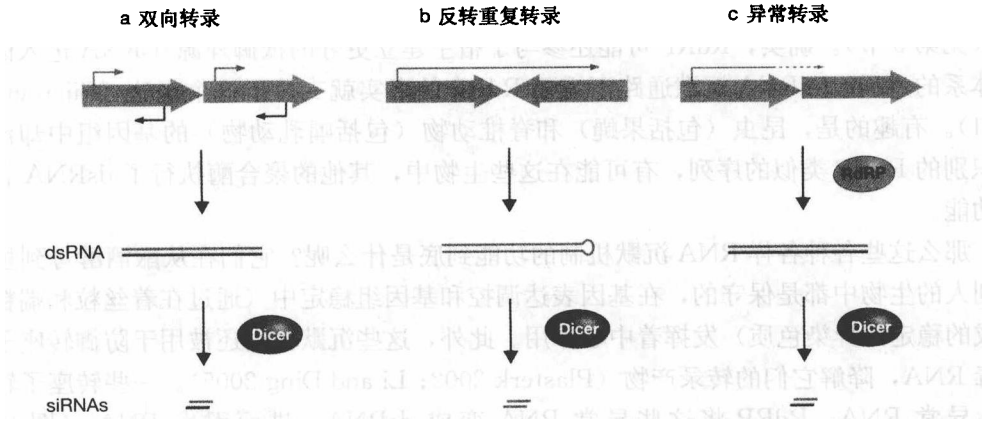



图 8-1　dsRNA 的来源，Dicer 核酸酶切割 dsRNA 产生 siRNA，siRNA 又是 RNA 沉默的扳机
(a) 在酿酒酵母着丝粒区重复序列区和接合型沉默位点的 *cenH* 区都有双向转录。(b) 许多植物和动物细胞中反向重复序列的转录可能产生 dsRNA。(c) 缺乏合适加工信号的非正常 RNA 的转录也可能引起 RdRP 合成 dsRNA。

到目前为止，已经鉴定出两个整合有 siRNA 的复合体：RISC 和 RITS。在 RNA 诱导沉默复合体（RISC）中，siRNA 通过碱基配对识别目标 mRNA，起始 siRNA 配对 mRNA 区域的内切降解（Hannon 2002；Bartel 2004)。Argonaute/PIWI 家族蛋白(RISC 的一个组分）的 RNaseH 结构域执行这个起始的 mRNA 切割事件。在核内 RNA 诱导转录沉默（RITS）复合体（类似于 RISC）中，siRNA 定位复合体到染色体的某些区域，引起相应的染色质修饰（Verdel et al. 2004；Buhler et al. 2006)。这就是

本章聚焦的 RITS 介导的 RNA 通路。

主要的 Argonaute 和 Dicer 蛋白还参与了另外一种涉及 microRNA（miRNA）的 RNA 沉默机制。内源非编码基因转录产生的 RNA，由于其展开的 dsRNA 区域而形成发卡 RNA 结构，再经过一系列步骤加工形成 miRNA（Bartel. 2004；Filipowicz et al. 2005）。和 siRNA 类似，miRNA 长 21～24nt，通过 Argonaute 蛋白形成部分 RISC 复合体，定位到特异的 mRNA 上。这种定位能够导致 PIWI/RNaseH 结构域对 mRNA 的剪切和结合 mRNA 5′端 7meG 帽子结构抑制翻译。这可能还与将 mRNA 运输到细胞质处理 RNA 的细胞器 P 小体（processing bodies，处理小体）的过程相偶联。因此，尽管 siRNA 和 miRNA 通过相似的机制来灭活其识别的 mRNA，但它们的产生却是通过至少两条不同的 dsRNA 加工途径。miRNA 通路之所以能够辨别自己，是因为所有的 miRNA 都是由内源非编码基因产生的，这些基因主要是受发育调控的，反过来，这些 miRNA 又主要定位在同源基因上并随着发育调控它们的沉默。

尽管 dsRNA 可以由双向转录产生的正义链和反义链复性生成或存在于发卡结构中，但在有些细胞中，RNAi 还需要另外一种酶来产生 dsRNA。这就是在植物和线虫中发现的 RNA 指导的 RNA 聚合酶（RNA-directed RNA polymerase，RdRP）（Dalmay et al. 2000；Sijen et al. 2001）。它用 siRNA 作为引物产生更多的 dsRNA，然后这些 dsRNA 再被 Dicer 加工成 siRNA。因此，RdRP 的主要功能就被认为是扩增 RNAi 效应，但是，就像后面要讨论的，RdRP 可能在 dsRNA 合成的起始中有着更多的特殊功能（见第 5 节）。确实，RdRP 可能还参与了宿主建立更好的抵御外源 dsRNA 进入的防御体系的过程。miRNA 沉默通路不需要 RdRP 的事实就支持了这个想法（Sijen et al. 2001）。有趣的是，昆虫（包括果蝇）和脊椎动物（包括哺乳动物）的基因组中却没有可识别的 RdRP 类似的序列，有可能在这些生物中，其他的聚合酶执行了 dsRNA 合成的功能。

那么这些各种各样 RNA 沉默机制的功能到底是什么呢？它们在从酿酒酵母到植物再到人的生物中都是保守的，在基因表达调控和基因组稳定中（通过在着丝粒和端粒区形成的稳定的异染色质）发挥着中心作用。此外，这些沉默机制还被用于防御转座子和病毒 RNA，降解它们的转录产物（Plasterk 2002；Li and Ding 2005）。一些转座子转录产生异常 RNA，RdRP 将这些异常 RNA 变成 dsRNA，进而引发 RNAi（图 8-1）（Baulcombe 2004）。

## 2. 关于 RNA 介导基因转录水平沉默的早期证据

酿酒酵母和拟南芥中有几个关于 RNAi 介导染色质修饰的已经研究地很透彻的例子。在讨论这几个例子之前，我们先简要讨论一下表明 RNA 介导染色质和 DNA 修饰的几个早期的实验。RNA 中间体在 TGS 中作用的最早证据来源于对植物类病毒的研究。马铃薯纺锤形块茎类病毒（potato spindle tuber viroid，PSTV）的基因组长 359nt，它通过 RNA-RNA 途径进行复制。PSTV 导入土豆基因组中导致核内同源序列的 DNA 甲基化（Wassenegger et al. 1994）。然而，这些和整合到基因组中的 PSTV DNA 拷贝

只在支持类病毒 RNA 转录的植物中被甲基化，这表明有 RNA 中间体指导了 DNA 甲基化（Wassenegger et al. 1994）。在拟南芥中，异常 RNA 的产生导致了所有同源启动子区的 DNA 甲基化和基因转录水平沉默（Mette et al. 1999）。这个结果，还有病毒基因组在植物中的复制生成的小 RNA 大小都在 22 nt 大小的结果都说明与 RNAi 相联系的机制介导了 DNA 甲基化（Mette et al. 2000）。这些实验结果，以及最早在矮牵牛和土豆中发现的重复序列转基因诱导基因沉默的现象，都被广泛地认为 RNAi 介导的基因沉默的最早期例子（Napoli et al. 1990；见第 9 章）。

RNAi 和 TGS 联系的进一步证据来自于对果蝇中重复序列诱导的基因沉默现象的研究（见第 5 章）。转基因多个串联拷贝的导入同时导致转基因和内源拷贝的沉默（Pal-Bhadra et al. 1999）。这个沉默需要 chromo 结构域蛋白 Polycomb 的参与，Polycomb 同时也参与了将同源异型调控基因包装成异染色质样结构的过程（Francis and Kingston 2001）。另外这个重复序列诱导的基因沉默还需要 Piwi 蛋白的参与，Piwi 蛋白是果蝇中 RNAi 所需要的 Argonaute 蛋白家族的成员（Pal-Bhadra et al. 2002）。在四膜虫中，小 RNA 的积累和原生动物体细胞大核中发现的大段 DNA 缺失需要另一个 Piwi 蛋白家族成员——Twil 的参与（见第 7 章）。这些结果以及在本章中讨论的更多的最新研究都表明，RNAi 通路参与了果蝇中抑制性染色质结构的组装。

其他重复序列诱导的沉默机制已在丝状真菌中被发现了，包括粗糙链孢霉菌中发现的重复序列诱导的点突变（repeat-induced point mutation，RIP）和 *Ascobolus immersus* 中发现的减数分裂前诱导的甲基化（methylation induced pre-meiotically，MIP）。这些似乎并不涉及 RNA 中间体，因为它们的发生并不依赖该位点的转录状态（Galagan and Selker 2004）。然而，RIP 和 MIP 涉及配对位点，（举例来说）基因三个拷贝中的两个被沉默了，表明某种涉及同源位点的 DNA-DNA 相互作用的机制诱导了基因沉默。相反，减数分裂中非配对基因（silencing of unpaired DNA in meiosis，MSUD）的沉默（在链孢霉菌中也同样发生）则涉及 RNAi 通路（Shiu et al. 2001；见第 6 章的讨论），在其他生物体中也有相应的机制，包括在线虫中（Maine et al. 2005；见第 15 章）。

## 3. 酵母中的 RNAi 和异染色质组装

酵母染色体中含有大量的异染色质 DNA 区域，主要是位于着丝粒区和接合型沉默位点（*mat2/3*）的重复 DNA 元件（Grewal 2000；Pidoux and Allshire 2004）。每个酿酒酵母的着丝粒都包含一个唯一的中心核心区（central core region，*cnt*），这个核心区被两种类型的重复序列包围着，这两种类型的重复序列被称为接合型沉默（silencing mating-type）最内（innermost，*imr*）和最外（outermost，*otr*）重复序列（图 8-2）。*otr* 区本身又是由 *dh* 和 *dg* 重复序列组成。酵母中异染色质的形成涉及许多反式作用因子的相关作用，包括组蛋白去乙酰化酶（histone Deacetylases，HDAC）、Clr4（一个组蛋白 H3K9 甲基转移酶，histone H3K9 methyltransferase，HKMT）、组蛋白 H3K9-methyl 结合蛋白、Swi6（HP1 同源蛋白）和 Chp1。最开始募集 Swi6 和 Clr4 到染色质被认为导致了 H3K9 甲基化的扩散和异染色质的形成，Clr4 催化了组蛋白甲基化，Swi6 通过

自身的相互作用介导了甲基化向邻近核小体的延伸（Grewal and Moazed 2003）。

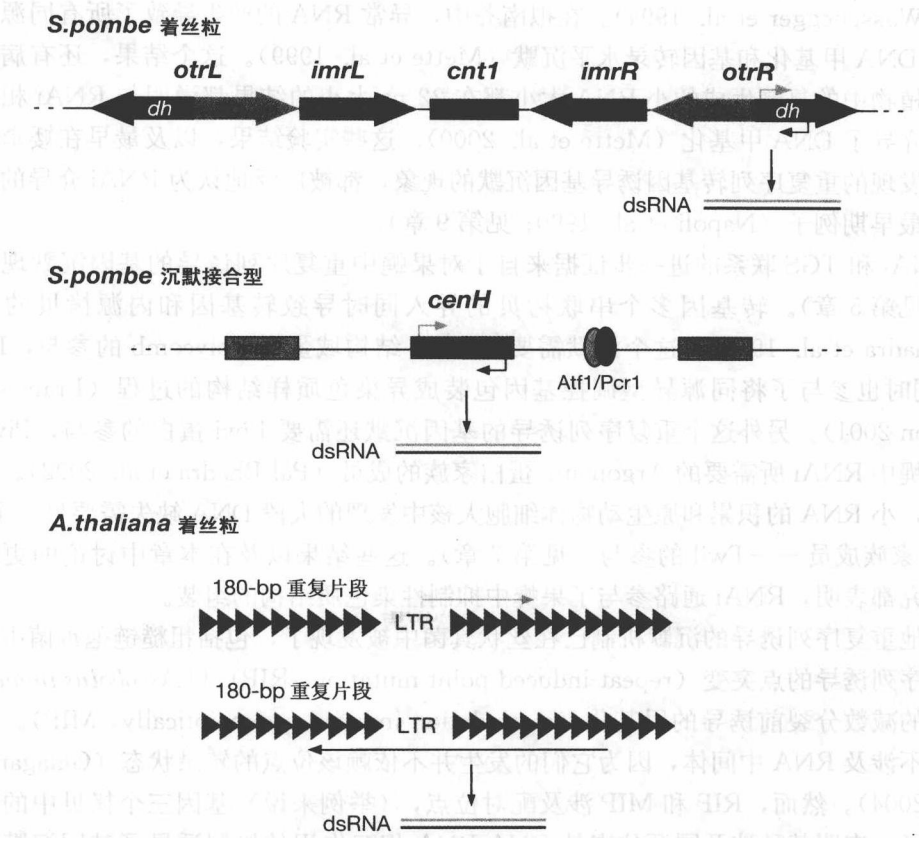



图 8-2 酵母和拟南芥中染色体异染色质区域的组织

酵母第 1 号染色体的着丝粒区如图所示。唯一的中心核心（*cnt1*）区被 *innermost*（*imrL* 和 *imrR*）和 *outermost*（*otrL* 和 *otrR*）重复序列围绕着。*otr* 区双向转录，产生正向（蓝色）和反向（红色）转录物。*mat2* 和 *mat3* 基因之间的区域包含一个与着丝粒区 dg 和 dh 重复序列（*cenH*）同源的区域，也是双向转录的。Atf1 和 Pcr1 是 DNA 结合蛋白，它们的作用与接合型沉默的 RNAi 相对应。拟南芥的着丝粒由 180bp 的重复序列（绿色）和散布其间的逆转座元件（黄色）组成。在逆转录因子长末端重复序列（long terminal repeat，LTR）内起始的正向转录物和 180bp 的重复序列内起始的反向转录物如图所示。

RNAi 通路中组分的突变惊人地导致了着丝粒异染色质的丢失和每个 *dg* 和 *dh* 重复元件内双向启动子非编码的正向和反向转录物的积累（图 8-2）（Volpe et al. 2002）。酿酒酵母含有编码每个 RNAi 通路蛋白的单一基因，Dicer、Argonaute 和 RdRP（分别是 *dcr1*$^+$、*ago1*$^+$和 *rdp1*$^+$）。这些基因的任何一个被删除后都会导致组蛋白 H3K9 甲基化的丢失，而且突变体表现出染色体分离缺陷，而这主要是与异染色质组装缺陷有关（Provost et al. 2002；Volpe et al. 2003）。还有，酿酒酵母小 RNA 库的测序鉴定出了特异定位在着丝粒重复序列区域和核糖体 DNA 重复序列上的约 22 nt 的 RNA，这意味着 cenRNA 可以产生能加工成 siRNA 的 dsRNA（Reinhart and Bartel 2002）。因此，这意味 RNAi 通路可以募集 Swi6 和 Clr4 到染色质，起始或维持在上述每个位点的异染色

质形成（图 8-3）（Hall et al. 2002；Volpe et al. 2002）。

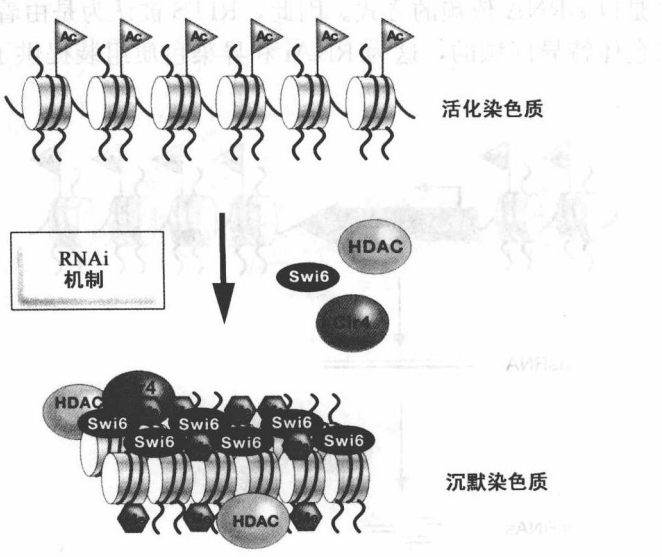



图 8-3　异染色质的组装

涉及组蛋白修饰酶（HDAC 和 Clr4）的相关作用和组蛋白结合蛋白（如 Swi6），并且能够由 RNAi 机器指导 HDAC 去乙酰化，Clr4 结合和组蛋白 H3K9 甲基化依次发生。Swi6 结合到 H3K9 甲基化的组蛋白尾巴上，伴随着 Swi6 寡聚化的 H3K9 甲基化的顺序循环导致了甲基化的蔓延。

有趣的是，TGS 和 PTGS 机制似乎都有助于 *cen*RNA 的下调。正向链转录物主要是在转录水平被沉默，RNAi 突变证明了这一点（Volpe et al. 2002）。然而，*cen* 转录物的反向链却不受 Swi6 突变的影响（Volpe et al. 2002），并且这个 *cen* 反向转录物的沉默主要发生在转录后水平。

RNAi 也在接合型位点（*mat2/3*）的沉默中起作用（Hall et al. 2002）。*mat2/3* 被与着丝粒区重复序列高度同源的一段 DNA（*cenH*，cenHomology）打断（图 8-2）。与 *cen* 重复序列一样，*cenH* 区也是双向转录的，产生正向和反向转录物（Noma et al. 2004）。在 RNAi 突变体中这些 *cenH* 转录物积累到很高的水平。相反，RNAi 不是插入到 *mat2/3* 转基因报告基因沉默所必需的，除非沉默不是一开始被无端的削弱。这种差别的原因在于部分酶基因的沉默机制涉及两个 DNA 结合蛋白——Pcr1 和 Atf1，这两个蛋白结合 *mat2/3*，不依赖 RNAi 地募集异染色质机器（Jia et al. 2004）。这足以在缺乏 RNAi 的情况下沉默报告基因了，但并不能阻止非编码 *cenH* 转录物的积累（图 8-2）。

## 4. 小 RNA 与 RNAi 效应蛋白复合体共同起始异染色质组装

已经发现酵母中异染色质的形成和其他系统中转录基因沉默涉及 RNAi 通路，而这又引起一个问题：RNAi 是如何直接调控染色质结构的呢？Chp1 是一种 chromo 结构域蛋白，是异染色质的结构组分。纯化 Chp1 发现了 RITS 复合体（Verdel et al. 2004）。RITS 除 Chp1 之外，还包含酿酒酵母的 Ago1 蛋白和未知功能的 Tas3 蛋白。RITS 复

合体也含有 Dicer 核酸酶切割产生的着丝粒区 siRNA，非常重要的是，RITS 结合在着丝粒重复序列区是以 siRNA 依赖的方式。因此，RITS 被认为是由着丝粒区 siRNA 引导定位并灭活染色体特异区域的，这为 RNAi 和异染色质组装提供了一个直接的联系（图 8-4）。

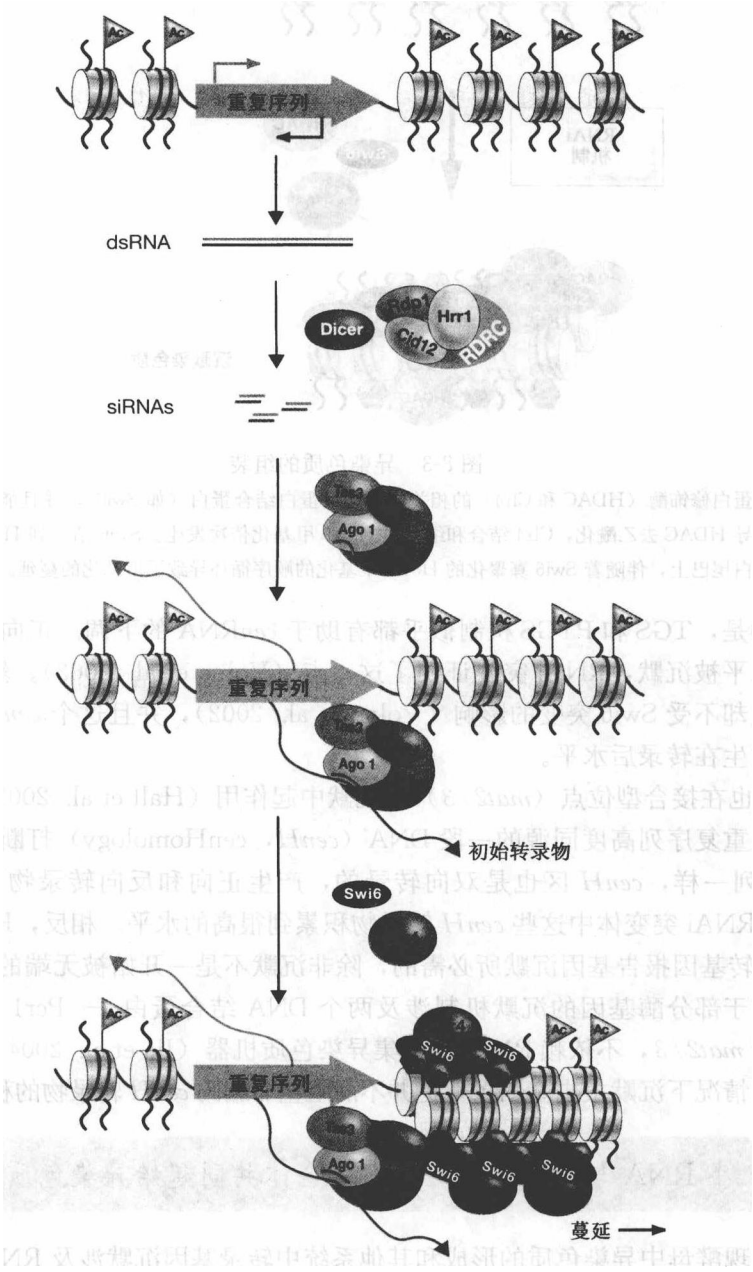



图 8-4 酵母中的 RNAi 和 siRNA 指导的异染色质的组装

siRNA 的产生既需要 Dicer 也需要 RDRC。定位的起始涉及 RITS 和 siRNA 介导的与相应转录物的识别。Chp1 的 chromo 结构域与 H3K9 甲基化的组蛋白结合稳定了 RITS 的结合。Clr4 和 Swi6 的募集介导了 H3K9 甲基化的蔓延。

和介导 PTGS 的 RISC 复合体一样，RITS 也用 siRNA 来识别定位的。与 RISC 不同的是，RITS 复合体结合在染色质并起始异染色质的形成而不是 mRNA 的失活。siRNA如何能够定位在特异的染色体区域？有两种可能的机制被提出来。在第一个模型里，与 RITS 复合体中 Ago1 结合的 siRNA 与展开的 DNA 双螺旋配对结合。在第二个模型里，RITS 结合的 siRNA 与目标位点非编码 RNA 转录物配对结合（图 8-4）。

在这两个模型中，RITS 通过 siRNA 结合到染色质，募集 Clr4 HKMT 并导致随后的 H3K9 甲基化。Swi6 随后结合上来并导致 H3K9 甲基化的扩散和异染色质的形成。然而，RITS 与染色质的结合也需要 Clr4，Clr4 提供了 RITS 复合体能结合的甲基化 H3K9，因此稳定了 RITS 与染色质的结合。已经知道 Chp1 的 chromo 结构域特异地结合甲基化的 H3K9 残基（Partridge et al. 2002）。这个相互作用中涉及的 Clr4 和 Chp1 chromo 结构域的突变都导致了 RITS 与染色质结合的丢失（Partridge et al. 2002；Noma et al. 2004）。另外，RITS 也能在缺乏 siRNA 的情况下，通过 Chp1 的 chromo 结构域在 *mat2/3* 和端粒区与被甲基化 H3K9 包被的染色质区域结合（Noma et al. 2004；Petrie et al. 2005）。总之，RITS 复合体能结合染色质，通过 Chp1 与甲基化 H3K9 的结合和 siRNA 与 DNA 或 RNA 转录物的配对结合。

最新的证据有力地支持了 RITS 和 siRNA 在异染色质组装起始中地作用。Buhler 等（2006）用一个位点特异的 RNA 结合蛋白人工地将 RITS 复合体聚集到正常情况下活跃的 *ura4*$^+$基因的 RNA 转录物上。这种聚集显著地导致了 *ura4*$^+$ siRNA 的产生和 *ura4*$^+$基因的沉默，这种作用同时依赖 RNAi 和异染色质组分。另外，这个系统有利于对新产生的 siRNA 起始 H3K9 甲基化和 Swi6 结合的能力做出估测，而它们又是异染色质形成的分子标记。有趣的是，新产生的 *ura4*$^+$siRNA 受保守的 siRNA 核酸酶 Eri1 负向调控，Eri1 将 siRNA 限制在它们产生的地方。然而，当编码 Eri1 的基因被删除后，*ura4*$^+$siRNA 能反式作用沉默 *ura4*$^+$基因的第二个拷贝，这个拷贝在同一个细胞中被插入在另外一条染色体上。这个实验证明 siRNA 可以作为一个特殊因子定位 RITS 和异染色质组装到基因组中曾经活跃的区域。

酵母中 siRNA 起始基因沉默的能力被用另外一种方法检查过，这个方法依赖于能产生与 GFP 转基因同源的 siRNA 的发卡 RNA 的表达（Sigova et al. 2004）。在这个系统中，发卡 siRNA 促进了 GFP 报告基因在 PTGS 水平的沉默，而不是 TGS 水平的（Sigova et al. 2004）。现在还不清楚为什么发卡 siRNA 不能诱导染色体上 GFP 拷贝的 TGS 和异染色质组装。一个可能的解释是，异染色质在特异的核内亚区域组装，在这些区域之外并不能有效地组装（Gasser et al. 2004；见第 4 章）。

### 5. dsRNA 的合成和 siRNA 的产生

酵母中着丝粒 DNA 重复元件的双向转录原则上可以作为 dsRNA 的起源（Volpe et al. 2002）。正向和反向转录物退火产生的 dsRNA 可作为 Dicer 核酸酶的底物。然而，RNA 指导的 RNA 聚合酶和它的共作用因子，以及 Clr4 HKMT 都参与了由 Dicer 剪切产生 siRNA 的过程（Hong et al. 2005；Li et al. 2005；Buhler et al. 2006）。这些结果

表明由 Dicer 剪切产生的异染色质 siRNA 与染色质和 Rdp1 依赖的事件相联系（图 8-4）。

Rdp1 酶位于一个多蛋白的复合体中，这个复合体还包括 Hrr1、RNA 螺旋酶、Cid12-DNA 聚合酶 beta 家族的成员，poly（A）聚合酶也属于这个家族（Motamedi et al. 2004）。这个复合体被命名为 RNA 指导的 RNA 聚合酶复合体（RNA-directed RNA polymerase complex，RDRC），它所有的亚基都参与了在着丝粒区形成异染色质的过程（Motamedi et al. 2004）。就像从有 Rdp1 存在所推断的一样，RDRC 在体外具有 RNA 指导的 RNA 聚合酶活性，导致这种活性丧失的突变体同样在体内导致了 RNAi 依赖的基因沉默的丢失（Motamedi et al. 2004；Sugiyama et al. 2005）。RDRC 体外 RNA 合成的活性不需要 siRNA 引物（Motamedi et al. 2004）。RITS 可能通过募集 RDRC 提供体内的特异性，RDRC 通过 siRNA 选择模板。与这个假说一致的是，RDRC 的亚基参与了 siRNA 的产生，从细胞内纯化的不含有 RDRC 任何亚基的 RITS 复合体中没有 siRNA（Motamedi et al. 2004；Li et al. 2005；Sugiyama et al. 2005；Buhler et al. 2006）。Cid12 在 RDRC 中的存在是有趣的，同时引出了一种可能性，另外一种聚合酶活性参与了染色体相关的 RNA 沉默。由于这个家族的一些成员具有 poly（A）聚合酶的活性，所以有一种可能就是，Rdp1 产生的 dsRNA 的腺嘌呤化功能对其进一步加工是重要的。有趣的是，Cid12 样蛋白在整个真核生物界都是保守的（表 8-1）；Rde-3 是线虫中这个蛋白家族的成员，它的突变导致了 RNAi 的缺陷（Chen et al. 2005），证实了这些在 RNAi 通路中的保守性。

有相关的证据表明，dsRNA 的合成和加工与染色体上非编码着丝粒 RNA 转录位点上发生的异染色质 siRNA 产生有关（图 8-4）。这些证据是：第一，Rdp1 可以和着丝粒 DNA 重复序列交联（Volpe et al. 2002；Sugiyama et al. 2005），也可以和这些区域产生正反向转录物交联（Motamedi et al. 2004）。与同 DNA 的交联一样，与着丝粒 RNA 的交连也需要 Dicer 和 Clr4，是 siRNA 和染色质依赖的。第二，siRNA 的产生需要染色质组分，包括 Clr4、Swi6 和 HDAC Sir2（Hong et al. 2005；Li et al. 2005；Buhler et al. 2006）。最后，RDRC 和 RITS 的结合依赖 siRNA 和 Clr4，表明发生在染色质上（Motamedi et al. 2004）。因此，dsRNA 和异染色质 siRNA 的产生可能涉及募集 RDRC 到与染色质相连的新生 pre-mRNA 转录物，如图 8-5 所示（Martienssen et al. 2005；Verdel and Moazed 2005）。转录和 siRNA 的产生很可能是同时发生的事实强调了 RNA 沉默机制间的差别，这些机制介导了染色质修饰和 PTGS。然而，这个差别也不可能是绝对的。举个例子，在线虫中，几个染色质组分的突变会导致和酵母中类似的 RNAi 和转座子诱导的 RNA 沉默的缺陷（见表 8-1）（Sijen and Plasterk 2003；Grishok et al. 2005；Kim et al. 2005），引出一种可能，不管沉默发生在 TGS 还是 PTGS 水平上，有些 dsRNA 的合成和加工发生在染色体上。

**表 8-1 RNAi 和异染色质蛋白的保守性**

| *S. pombe* | *A. thaliana* | *C. elegans* | *Drosophila* | *H. sapiens* |
|---|---|---|---|---|
| Dcr1 | DCL1～4 | Dcr-1 | Dcr1 和 2 | DCR-1 |
| Ago1 | AGO1～10 | Rde-1，Alg-1 和-2 | Ago1～3，Piwi | AGO-1～AGO-4 |
| | PRG-1、2 和 19 个其他的 | Aubergine/Sting | Piwi-1～Piwi-4 | PIWI-1～PIWI-4 |

续表

| *S. pombe* | *A. thaliana* | *C. elegans* | *Drosophila* | *H. sapiens* |
|---|---|---|---|---|
| Chp1[a] | CMT3 | — | — | — |
| Tas3[b] | — | — | — | — |
| Rdp1 | RDR1～6 | Ego-1，Rrf-1 to -3 | — | — |
| Hrr1 | SGS2/SDE3 | ZK1067.2 | GH20028p | KIAA1404 |
| Cid12 | —[c] | Rde-3，Trf-4[c] | CG11265[c] | POLS[c] |
| Swi6 | LHP1(TFL2) | Hpl-1，Hpl-2，F32E10.6 | HP1 | HP1α，β，γ |
| Clr4 | SUVH2～6 | —[d] | Su(var)3-9 | SUV39H1 和 2 |
| Rik1[e] | DDB1 | M18.5 | Ddb1 | DDB1 |
| Cul4 | CUL4 | Cul4 | Cul4 | CUL4 |
| Sir2 | SIR2 | Sir2-1 | Sir2 | SIRT1 |
| Eri1 | ERI1 | Eri-1 | CG6393 | THEX1 |

a chromo 结构域蛋白 Chp1 在表中所列的其他模型生物中没有发现明显的正同源物，但是大部分真核细胞都含有多种 chromo 结构域蛋白。CMT3 是拟南芥中 chromo 结构域 DNA 甲基转移酶，和 AGO4 在同一个通路中起作用，可能与 Chp1 同源。

b Tas3 没有明显的同源物被发现，但与鼠卵巢睾丸特异蛋白（NP_035152）有着微弱的序列相似性。

c Cid12 属于一个保守的蛋白大家族，它们与经典的 poly（A）聚合酶以及 2′-5′-寡聚腺嘌呤核苷酸酶有着序列相似性。

d 在线虫中有大约 20 种 SET 结构域蛋白，但是线虫中还没有发现任何一个 H3K9 HKMT。

e 酵母中有另一个 Rik1 样的蛋白，Ddb1，它参与了 DNA 损伤修复。多细胞动物和植物中似乎只有一个 Rik1 样的基因，称为 Ddb1，其被证明参与了 DNA 损伤修复，但是还不知道它是否参与了异染色质形成。

## 6. RNA-RNA 和 RNA-DNA 识别模型

要弄清 RNAi 在异染色质组装中作用的一个突出问题是，RITS/RDRC 是否与 DNA 或新生 RNA 结合？RNAi 机器组分聚集到基因转录物上能顺式诱导异染色质依赖的基因沉默的研究结果清楚地表明，这个过程能够被通过与新生 RNA 转录物的相互作用而加强（Buhler et al. 2006）。重要的是，顺式限制排除了 dsRNA 合成和 siRNA 产生的起始事件发生在成熟转录物上的可能性，因为不同等位基因的 mRNA 产物是不能被区分开的。还有一个 RNA-RNA 相互作用模型的直接推测，目标位点的转录参与了 RNAi 介导的异染色质组装。尽管需要转录这一点还没有被直接检测过，但是 RNA 聚合酶 II 两个不同亚基 Rpb2 和 Rpb7 上的突变可导致 siRNA 的生成和异染色质组装的特异缺陷，而不是总转录的缺陷（Djupedal et al. 2005；Kato et al. 2005）。这使人联想到 Rbp1 突变体。Rbp1 突变体在转录延伸相关的组蛋白修饰（即 H3K4 甲基化和 H2B 泛素化）上有缺陷（Hampsey and Reiberg 2003），为 RNAi 介导的 H3K9 甲基化和异染色质形成可以通过 RNAi 复合体与 RNA pol Ⅱ的结合而与转录延伸联系在一起的假说提供了例证。事实上，与获得广泛承认的异染色质是不可接近的结构并抑制转录的观点相矛盾的是，RNAi 介导的异染色质形成对 RNA pol Ⅱ与酵母着丝粒重复序列的结合几乎没有作用（Volpe et al. 2002；Djupedal et al. 2005；Kato et al. 2005；Buhler et al. 2006）。因此，新生 RNA 转录物，在 RNA-RNA 识别模型中以 RITS 模板起作用，存

在于异染色质区域（图 8-4）。

通过体内交联实验证明 RITS 和 RDRC 的组分可定位在非编码的着丝粒区 RNA 上，这个实验结果也支持 RNA-RNA 定位模型（Motamedi et al. 2004）。这种定位依赖于 siRNA，这意味着它涉及与非编码 RNA 碱基配对相互作用。另外，这还需要 Clr4 HKMT 的参与，意味着与 RITS 和甲基化 H3K9 相结合且发生在染色质上。不过，siRNA 可以直接通过碱基配对相互作用识别 DNA 的可能不能被排除。举个例子，在植物中，与不转录的启动子区互补的 siRNA 还能指导 DNA 甲基化，这是这些区域异染色质形成中的另一个修饰（见第 9 章）。

## 7. RNAi 如何募集染色质修饰酶？

募集 Clr4 和 Swi6 是启动组蛋白 H3K9 甲基化和异染色质组装的一个关键步骤，它可能通过一种自调节修饰的模式（图 8-3，图 8-4）（Grewal and Moazed 2003）。然而，由于 RITS 结合染色质和 Clr4 催化组蛋白 H3K9 甲基化是相互依赖的过程，所以很难判断哪个事件提供了依赖 RNAi 的异染色质组装的引子。这个“先有鸡还是先有蛋”问题的一个答案是，siRNA 依赖的碱基配对相互作用可能提供了异染色质组装的起始信号（图 8-4）。与这个假说一致的是，*ura4*$^+$ siRNA 的产生促进了 *ura4*$^+$ 基因曾经活跃拷贝的沉默，并与募集 RITS 和 Swi6 到染色质相联系（Buhler et al. 2006）。然而，RITS 的起始结合也许是瞬时的，是很难检测到的；RITS 与染色质的紧密结合需要两个事件的相互作用：①RITS 结合的 siRNA 和新生转录物的相互作用和②Chp1 的 chromo 结构域结合到甲基化的 H3K9 上。在这个模型中，RITS 本身直接募集 Clr4，或者 Clr4 被一种相对应的通路所募集，这个通路涉及一个或更多的 DNA 结合蛋白，就像在接合型沉默和端粒区的例子一样（Jia et al. 2004；Kanoh et al. 2005）。在另外一个版本中，Clr4 介导的 H3K9 甲基化可能参与稳定 RITS 和染色质的结合，进而导致 RDRC 的募集、dsRNA 的合成和 siRNA 的形成（图 8-5）。

Clr4 最近被发现是一个多蛋白复合体的组分，这个复合体含有异染色质蛋白 Rik1、Cullin E3 泛素连接酶、Cul4 和几个其他蛋白（Hong et al. 2005；Horn et al. 2005；Jia et al. 2005；Li et al. 2005）。这些 Clr4 结合蛋白进一步强化了 RNA 和异染色质形成之间的联系。Rik1 是 beta propeller WD 重复蛋白大家族的成员，参与 RNA 和 DNA 的结合。这个家族的成员包括，涉及前体 mRNA 剪切的剪切多聚腺苷酸特异因子 A（cleavage polyadenylation specificity factor A，CPSF-A）和涉及结合 UV 损伤 DNA 的结合蛋白 1（DNA damage binding 1，Dbd1）。CPSF-A 有着特殊的意义，因为 Rik1 与其涉及 mRNA 多聚腺苷酸 RNA 推测的 RNA 结合结构域有着序列相似性（Barabino et al. 2000）。Dbd1 蛋白和 Rik1 一样是 Cul4 E3 泛素连接酶复合体的组分，参与了 UV 损伤 DNA 的识别和修复（Higa et al. 2003；Zhong et al. 2003）。一个令人兴奋的可能性是，Rik1 和 CPSF-A 和 Dbd1 以类似的方式在异染色质组装过程中结合于 RNAi 的产物上（图 8-5）。

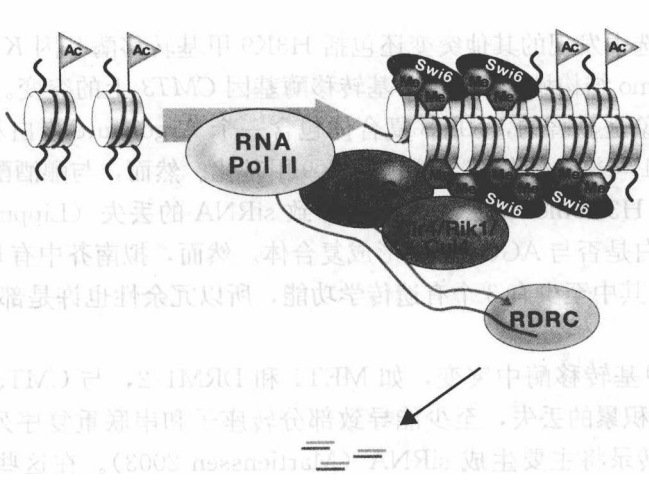



图 8-5　酵母中共转录 dsRNA 和 siRNA 产生和 Clr4-Rik1-Cul4 组蛋白甲基转移酶募集的模型

## 8. 拟南芥中 RNAi 介导的染色质和 DNA 修饰

植物中 RNAi 指导异染色质修饰的机制与酿酒酵母中机制类似，但仍有很多差异。最重要的差别是植物在许多抑制的异染色质区域有甲基化的 DNA；从这个方面来说，是与脊椎动物相似，而与线虫和果蝇不同（Lippman and Martienssen 2004）。在寻找可降低 RNA 介导的 TGS 表型的 4 个突变筛选中，发现了 H3K9 特异的 HKMT 和 RNAi 组分的突变，并揭示它也需要 DNA 甲基转移酶，SWI/SNF 重塑复合体和一个新的 RNA 聚合酶的作用（Baulcombe 2004）。这些筛选在第 9 章中有详细阐述，这里我们简要比较一下酿酒酵母和植物中的机制。

沉默突变体的每个筛选都用反式导入的反向重复序列来诱导内源或转基因报告基因的沉默。沉默的减轻意味着有一个突变发生在沉默通路的一个必需组分中。所用的内源基因是 *PAI2*（参与氨基酸的生物合成）（Mathieu and Bender 2004）和 *SUPERMAN*（一个调控花发育的转录因子）（Chan et al. 2004），转基因报告基因是由病毒或种子特异的强启动子驱动的（Matzke et al. 2004）。在每个例子中，沉默都发生在启动子上，在有些例子中还发生在基因的其他部分。这些筛选中发现的大量基因如图 8-6 所示（另见表 9-1）。在这些筛选中只鉴定出一个 RNAi 突变，即是 argonaute 基因 *AGO4*。其中的 3 个筛选获得了发生在 DNA 甲基转移酶中的突变，包括 MET1 和 CMT3。与哺乳动物 DNMT3 相关的第 3 个 DNA 甲基转移酶通过反向遗传被鉴定出来，这是由于其活性被 DRM1 和 DRM2 基因编码，而两个重复功能的基因无法在单突变筛选中被确定（见第 9 章）。确实，功能冗余性也许解释了筛选 RNAi 其他组分的失败，例如，尽管类 Dicer 酶（Dicer-like 3，DCL3）和 RNA 依赖的 RNA 聚合酶Ⅱ（RNA-dependent polymerase 2，RdRP）是转座子和和重复序列相关的 24nt siRNA 生成所必需的，在拟南芥中至少有两个其他的 *DCL* 基因能够在某种程度上替代 *DCL3*（Gasciolli et al. 2005）。

在 *PAI2* 筛选中发现的其他突变还包括 H3K9 甲基转移酶基因 *KYP/SUVH4* 中的突变和含有 chromo 结构域的 DNA 甲基转移酶基因 *CMT3* 上的突变。这个例子中与酿酒酵母对应的现象是显著的，RITS 复合体包含一个 Argonaute 蛋白和 chromo 结构域蛋白 Chp1、Chp1 与染色体的结合依赖 H3K9 甲基化。然而，与酿酒酵母不同的是，拟南芥中 CMT3 或 H3K9me2 的丢失并不会导致 siRNA 的丢失（Lippman et al. 2003），还不清楚这些蛋白是否与 AGO4 一起形成复合体。然而，拟南芥中有几个其他的 H3K9 特异的 HKMT，其中至少有 3 个有遗传学功能，所以冗余性也许是部分解释（Ebbs et al. 2005）。

其他 DNA 甲基转移酶中突变，如 MET1 和 DRM1/2，与 CMT3 中的突变相反，能够导致 siRNA 积累的丢失，至少能导致部分转座子和串联重复序列来源的 RNA 的积累，这些如果转录将主要生成 siRNA（Martienssen 2003）。在这些例子中，siRNA 的丢失与 H3K9me2 的丢失相关（Cao et al. 2003；Lippman et al. 2003）。SWI2/SNF2 染色质重塑 ATPase DDM1（降低 DNA 甲基化）的突变也驱除了 siRNA 和 H3K9me2 在大范围转座子上的积累，尽管当 siRNA 保留时，H3K9me2 也会保留（Lippman et al. 2003）。因此，在植物中 siRNA 通过结合甲基化 DNA 而不像在酵母中通过结合甲基化组蛋白来结合染色体（图 8-7）。

DDM1 还对转座子和重复序列有一个精确特异性，通过与基因的差别来识别它们。siRNA 也许是通过甲基结合蛋白结合到染色体上，它应该有其特异性来区别这种差异。与此一致的是，拟南芥中的转座子和重复序列是 24nt 和部分 21nt siRNA 的主要来源（Lippman et al. 2004）。着丝粒卫星重复序列，在每个着丝粒两侧都以数以万计的串联拷贝排列，它们也被转录和被 RNAi 加工（图 8-6）。这个加工依赖 DCL3、RDR2 和 DDM1。沉默还依赖 H3K9me2 和 CMT3。沉默比酿酒酵母中的复杂的多，逆转座子插入到重复序列能被沉默，这是依赖于包含有 MET1、DDM1 和组蛋白去乙酰化酶 HDA6 的其他机制（May et al. 2005）。

前面提到，酿酒酵母中 RNA 聚合酶Ⅱ的亚基参与沉默和 siRNA 生成，这支持了 RNAi 和染色质修饰装置被新生转录物募集到染色体上的假说（图 8-7）。在拟南芥中，一种新 RNA 聚合酶（RNA 聚合酶 Ⅳ）的两个亚基在上面提到的 4 个筛选中被找到（Kanno et al. 2005），但却是与 RNA 依赖的 RNA 聚合酶中的突变一起被第一次作为 PTGS 中的弱突变体被分离出来，（Herr et al. 2005）。还不清楚是什么模板被 RNA 聚合酶Ⅳ使用，甲基化 DNA（Onodera et al. 2005）和双链 RNA（Vaughn and Martienssen 2005）都可能是。RNA 聚合酶Ⅳ中只有最大的亚基是独特的，小亚基和 RNA 聚合酶Ⅱ中使用的一样。这些筛选中获得的其他 SWI2/SNF2 染色质重塑子可能改变了局部染色质结构促进了 RNA 聚合酶的活性。因此，很有可能是通过 RNA 聚合酶Ⅳ促进转录的（Kanno et al. 2004）。DDM1 可能也有类似的作用，转座子沉默需要 DDM1（也是一个染色质重塑子）远比需要 RNA 聚合酶Ⅳ或其他 SWI2/SNF2 蛋白更迫切。

基因能够被附近的转座子表观遗传沉默，拟南芥中的一个重要的例子就是印记的同源异型基因 *FWA*（Kinoshita et al. 2004）。这个基因的前两个外显子是非编码的并形成一个串联重复，这是因为 *SINE* 元件早已整合到这个位点上（Lippman et al. 2004）。

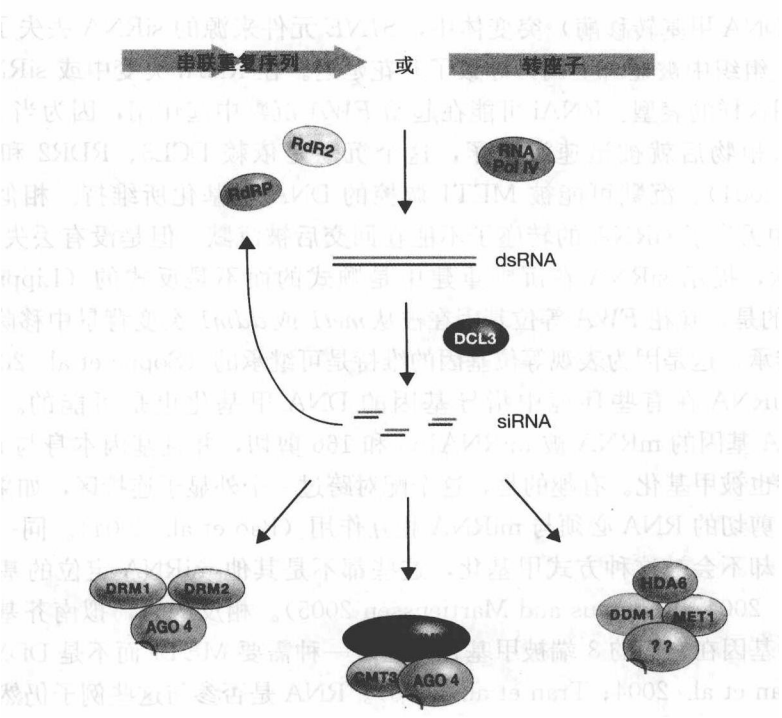



图 8-6　拟南芥中 RNAi 和 RNAi 介导的 DNA 和组蛋白甲基化必需的染色质蛋白总结
从重复 DNA 元件合成的 dsRNA 为 Dicer 介导的剪切和 siRNA 的产生（DCL3 和其他 Dicer）提供了底物。RNA 指导的 RNA 聚合酶（RNA-directed RNA polymerase，RdRP，RdR2）和 RNA 聚合酶 IV（RNA pol IV）可能直接参与了 dsRNA 的合成或扩增。siRNA 装载到 Argonaute 蛋白（如 AGO4）上，它很有可能与其他因子一起帮助定位到识别的重复序列，其 DNA 和 H3K9 被甲基化。

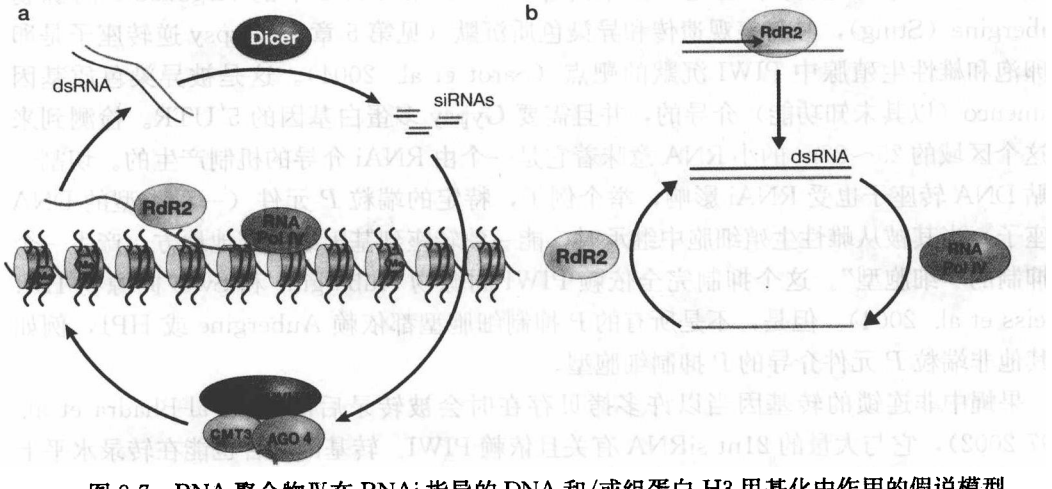



图 8-7　RNA 聚合物Ⅳ在 RNAi 指导的 DNA 和/或组蛋白 H3 甲基化中作用的假说模型
(a) RNA 聚合物Ⅳ转录甲基化的 DNA；RdR2 用 RNA 聚合物Ⅳ产物合成 dsRNA。然后 siRNA 指导甲基转移酶复合体结合到染色体上。(b) RNA 聚合物Ⅳ用 RdR2 合成的 dsRNA 模板来产生更多的 ssRNA 模板。

在 *met1*（即 DNA 甲基转移酶）突变体中，*SINE* 元件来源的 siRNA 丢失了，并且基因表达在分生组织中被显著上调，导致了开花延迟。在 RNAi 突变中或 siRNA 丢失时都没有观察到这样的表型。RNAi 可能在起始 *FWA* 沉默中起作用，因为当 *FWA* 转基因被首次导入植物后就被迅速沉默了，这个沉默是依赖 DCL3、RDR2 和 AGO4 的（Chan et al. 2004）。沉默可能被 MET1 调控的 DNA 甲基化所维持。相似的是，在 *met1* 突变体中丢失了 siRNA 的转座子不能在回交后被沉默，但是没有丢失 siRNA 的可以被再沉默，提示 siRNA 在沉默重建中是顺式的而不是反式的（Lippman et al. 2003）。类似的是，晚花 *FWA* 等位基因在被从 *met1* 或 *ddm1* 突变背景中移除后的回交中能稳定地继承，这是因为表观等位基因的维持是可继承的（Soppe et al. 2000）。

最后，miRNA 在有些环境中指导基因的 DNA 甲基化也是可能的。拟南芥中 *PHABULOSA* 基因的 mRNA 被 miRNA165 和 166 剪切，并且基因本身与 miRNA 配对的区域下游也被甲基化。有趣的是，这个配对跨过一个外显子连接区，如果这指导了甲基化的话，剪切的 RNA 必须与 miRNA 相互作用（Bao et al. 2004）。同一个基因家族的其他成员却不会被这种方式甲基化，这些都不是其他 miRNA 定位的基因（Martienssen et al. 2004；Ronemus and Martienssen 2005）。相反的是，拟南芥基因组中几个其他的典型基因在它们的 3′端被甲基化，通过一种需要 MET1 而不是 DDM1 的分子机制（Lippman et al. 2004；Tran et al. 2005）。RNA 是否参与这些例子仍然未知。

## 9. RNAi 介导的染色质修饰在动物中的保守性

表观遗传沉默研究最广泛的例子是在动物中发现的，包括果蝇、线虫和小鼠。RNA 和 RNA 干扰在转录沉默和异染色质修饰中的作用在有些模式动物中似乎是保守的，在原生动物和植物中也是。在果蝇中，PIWI 和 PIWI 中的 Argonaute 同源物 Aubergine（Sting），参与表观遗传和异染色质沉默（见第 5 章）。Gypsy 逆转座子是卵巢卵泡和雄性生殖腺中 PIWI 沉默的靶点（Sarot et al. 2004）。这是被异染色质基因 Flamenco（以其未知功能）介导的，并且需要 Gypsy 多蛋白基因的 5′UTR。检测到来自这个区域的 25～27nt 的小 RNA 意味着它是一个由 RNAi 介导的机制产生的。切割—粘贴 DNA 转座子也受 RNAi 影响，举个例子，特定的端粒 *P* 元件（一种类型的 DNA 转座子）当其被从雌性生殖细胞中继承时，能一致转座到基因组的其他地方，产生一个强抑制的“细胞型”。这个抑制完全依赖 PIWI 同源物 Aubergine 和 Swi6 同源物 HP1（Reiss et al. 2004）。但是，不是所有的 P 抑制细胞型都依赖 Aubergine 或 HP1，例如被其他非端粒 *P* 元件介导的 P 抑制细胞型。

果蝇中非连锁的转基因当以许多拷贝存在时会被转录后沉默（Pal-Bhadra et al. 1997 2002），它与大量的 21nt siRNA 有关且依赖 PIWI。转基因融合也能在转录水平上沉默彼此，但需要 Polycomb 染色质抑制子。这种沉默与转基因转录物来源的 siRNA 水平的增加无关，但很大程度上依赖 PIWI。在 PIWI 依赖的沉默中，像这个例子需要 Polycomb，其他则需要 HP1，或涉及 RNAi 通路和组蛋白甲基化。果蝇中转基因串联排列也表现出位置效应花斑，而且这个花斑效应被 *HP1* 的突变以及 *piwi*、*aubergine*

和推测的 RNA 螺旋酶 *Spindle-E*（*homeless*）的突变强烈抑制（Pal-Bhadra et al. 2004）。插入到着丝粒异染色质中的转基因也受影响，并且在 *spindle-E* 突变的细胞中，异染色质 H3K9me2 水平显著降低。这些结果有力支持了染色质蛋白和 RNAi 通路组分在果蝇异染色质内基因沉默中的作用。

在果蝇雄性生殖细胞中，异染色质 *Stellate* 抑制子重复序列［Su（ste）］位于 Y 染色体上，首先在反义链上转录，然后才在精母细胞发育中的两条链同时转录，这可能是在一个邻近转座子插入之后发生的（Aravin et al. 2001）。这些核转录物参与了紧密联系的 X 连锁的 *Stellate* 基因正义转录物的沉默，而 *Stellate* 基因的过表达可以导致精子发生的缺陷。尽管涉及异染色质序列，但这个例子中的沉默似乎是转录后的，与 25～27nt siRNA 相关，依赖于 *Aubergine* 和 *Spindle-E*。

在线虫中，体细胞中 TGS 的例子也有报道。这依赖 RNAi 通路基因 *rde-1*、*dcr-1*、*rde-4* 和 *rrf-1*，以及 HP1 同源物和组蛋白修饰装置（Grishok et al. 2005）。线虫中体细胞异染色质不是大范围延伸的，在生殖细胞中还是有自然发生的 RNAi 依赖的异染色质沉默的例子（Sijen and Plasterk 2003）。在减数分裂期间，未配对的序列（如雄性中的 X 染色体）被通过 H3K9me2 方式沉默。这种沉默依赖于 RNA 依赖的 RNA 聚合酶（Maine et al. 2005；见第 15 章），近似于红色链孢霉菌中的减数分离期未配对 DNA（meiotic silencing of unpaired DNA，MSUD）的沉默（见 Shiu et al. 2001；第 6 章）。然而，RNAi 装置中的其他组分并没有参与这个过程，也不清楚这与酿酒酵母中 RNAi 介导的异染色质组装是否有机制性联系。

最后，跟果蝇一样，哺乳动物细胞也缺乏与植物、线虫和真菌中发现的 RNA 依赖的 RNA 聚合酶有关的基因。反义 RNA 被用于大部分广泛研究的表观遗传现象，包括印记基因和 X 染色体失活（分别见第 19、17 章）。在 X 染色体失活的例子中，一个切割过的 17kb 多聚腺苷酸的非编码 RNA，被称为 *Xist*，参与表达它的 X 染色体的沉默。相反的，*Xist* 本身也在活跃的 X 染色体上被沉默，这个过程部分依赖于反义 RNA *Tsix*。沉默同时伴有上游染色质区域组蛋白的修饰改变，如 3K9me2 和 H3K27me3（见第 17 章）。小鼠中其他印记位点的沉默，包括 *Igf2r* 和 *Dlk1-Gtl2* 区域，也分别被来自父本等位基因和来自母本等位基因的反义转录物维持。在 *Dlk1-Gtl2* 例子中，这个非编码 RNA 特异地被加工成 miRNA，靶向来自母本等位基因的反义转录物。这个母本等位基因编码 *sushi*（*gypsy*）类的逆转座子（Davis et al. 2005）。

尽管有很多与酵母中来自异染色质重复序列的正向和反向转录物的相对应产物，但是 RNAi 本身在基因印记和 X 失活中的作用是独一无二的。不过，siRNA 导入癌细胞能引起染色质在同源启动子区被 H3K9me2 标记（Ting et al. 2005）。在有些例子中，小 RNA 通过结合 DNA 甲基转移酶和 DNA 甲基化结合蛋白（Jeffery and Nakielny 2004）还能导致 DNA 甲基化（Morris et al. 2004）。最后，Dicer 敲除的脊椎动物细胞系表现出染色体分裂缺陷，令人想到了在酿酒酵母突变体中发现的这种现象，伴随着异染色质形态上的变化和卫星重复序列的表达以及黏结蛋白的错误定位（Fukagawa et al. 2004；Kanellopoulou et al. 2005）。

## 10. 小 结

基因被小 RNA 分子调控的可能性早在 40 年前就已被提出（Jacob and Monod 1961），同时也提出了“控制 RNA”可能与重复序列有关（Britten and Davidson 1969）。自从在大肠杆菌和感染性噬菌体 λ 中鉴定出作为位点特异 DNA 结合蛋白的 lambda 和 *lac* 抑制子（Gillbert and Muller-Hill 1966；Ptashne 1967）后，基因调控的研究就几乎被全部聚焦在作为特殊因子的核酸结合蛋白上。小 RNA 分子作为特异因子在各种 RNA 沉默机制中作用的发现清楚地确定了 RNA 作为调控基因和它们表达产物的序列特异因子。酿酒酵母、拟南芥以及其他模式生物中的研究揭示了小 RNA 在介导基因组表观遗传修饰的显著作用，这些修饰调控基因沉默有助于建立基因组稳定所必需的异染色质区域和细胞核分裂。许多重要的机制性问题都还没有解决，未来的研究会给我们带来更多关于 RNA 如何调控基因表达的惊喜。

张道永 译
王大亮 校
孙方霖 审校

## 参 考 文 献


Aravin A.A., Naumova N.M., Tulin A.V., Vagin V.V., Rozovsky Y.M., and Gvozdev V.A. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr. Biol.* **11:** 1017–1027.

Bao N., Lye K.W., and Barton M.K. 2004. MicroRNA binding sites in *Arabidopsis* class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev. Cell* **7:** 653–662.

Barabino S.M., Ohnacker M., and Keller W. 2000. Distinct roles of two Yth1p domains in 3′-end cleavage and polyadenylation of yeast pre-mRNAs. *EMBO J.* **19:** 3778–3787.

Bartel D.P. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116:** 281–297.

Baulcombe D. 2004. RNA silencing in plants. *Nature* **431:** 356–363.

Bernstein E., Caudy A.A., Hammond S.M., and Hannon G.J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409:** 363–366.

Britten R.J. and Davidson E.H. 1969. Gene regulation for higher cells: A theory. *Science* **165:** 349–357.

Buhler M., Verdel A., and Moazed D. 2006. Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* **125:** 873–886.

Cao X., Aufsatz W., Zilberman D., Mette M.F., Huang M.S., Matzke M., and Jacobsen S.E. 2003. Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr. Biol.* **13:** 2212–2217.

Chan S.W., Zilberman D., Xie Z., Johansen L.K., Carrington, J.C., and Jacobsen S.E. 2004. RNA silencing genes control de novo DNA methylation. *Science* **303:** 1336.

Chen C.C., Simard M.J., Tabara H., Brownell D.R., McCollough J.A., and Mello C.C. 2005. A member of the polymerase β nucleotidyltransferase superfamily is required for RNA interference in *C. elegans*. *Curr. Biol.* **15:** 378–383.

Dalmay T., Hamilton A., Mueller E., and Baulcombe D.C. 2000. *Potato virus X* amplicons in *Arabidopsis* mediate genetic and epigenetic gene silencing. *Plant Cell* **12:** 369–379.

Davis E., Caiment F., Tordoir X., Cavaille J., Ferguson-Smith A., Cockett N., Georges M., and Charlier C. 2005. RNAi-mediated allelic *trans*-interaction at the imprinted *Rtl1/Peg11* locus. *Curr. Biol.* **15:** 743–749.

Djupedal I., Portoso M., Spahr H., Bonilla C., Gustafsson C.M., Allshire R.C., and Ekwall K. 2005. RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev.* **19:** 2301–2306.

Ebbs M.L., Bartee L., and Bender J. 2005. H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Mol. Cell. Biol.* **25:** 10507–10515.

Elbashir S.M., Lendeckel W., and Tuschl T. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15:** 188–200.

Filipowicz W., Jaskiewicz L., Kolb F.A., and Pillai R.S. 2005. Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr. Opin. Struct. Biol.* **15:** 331–341.

Fire A., Xu S., Montgomery M.K., Kostas S.A., Driver S.E., and Mello C.C. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391:** 806–811.

Francis N.J. and Kingston R.E. 2001. Mechanisms of transcriptional memory. *Nat. Rev. Mol. Biol.* **2:** 409–421.

Fukagawa T., Nogami M., Yoshikawa M., Ikeno M., Okazaki T., Takami Y., Nakayama T., and Oshimura M. 2004. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat. Cell Biol.* **6:** 784–791.

Galagan J.E. and Selker E.U. 2004. RIP: The evolutionary cost of genome defense. *Trends Genet.* **20:** 417–423.

Gasciolli V., Mallory A.C., Bartel D.P., and Vaucheret H. 2005. Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing *trans*-acting siRNAs. *Curr. Biol.* **15:** 1494–1500.

Gasser S.M., Hediger F., Taddei A., Neumann F.R., and Gartenberg M.R. 2004. The function of telomere clustering in yeast: The circe

effect. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 327–337.
Gilbert W. and Muller-Hill B. 1966. Isolation of the Lac repressor. *Proc. Natl. Acad. Sci.* **56:** 1891–1898.
Grewal S.I. 2000. Transcriptional silencing in fission yeast. *J. Cell. Physiol.* **184:** 311–318.
Grewal S.I. and Moazed D. 2003. Heterochromatin and epigenetic control of gene expression. *Science* **30:** 798–802.
Grishok A., Sinskey J.L., and Sharp P.A. 2005. Transcriptional silencing of a transgene by RNAi in the soma of *C. elegans*. *Genes Dev.* **19:** 683–696.
Hall I.M., Shankaranarayana G.D., Noma K., Ayoub N., Cohen A., and Grewal S.I. 2002. Establishment and maintenance of a heterochromatin domain. *Science* **297:** 2232–2237.
Hamilton A.J. and Baulcombe D.C. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286:** 950–952.
Hampsey M. and Reinberg D. 2003. Tails of intrigue: Phosphorylation of RNA polymerase II mediates histone methylation. *Cell* **113:** 429–432.
Hannon G.J. 2002. RNA interference. *Nature* **418:** 244–251.
Herr A.J., Jensen M.B., Dalmay T., and Baulcombe D.C. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308:** 118–120.
Higa L.A., Mihaylov I.S., Banks D.P., Zheng J., and Zhang H. 2003. Radiation-mediated proteolysis of CDT1 by CUL4-ROC1 and CSN complexes constitutes a new checkpoint. *Nat. Cell Biol.* **5:** 1008–1015.
Hong E.E., Villen J., Gerace E.L., Gygi, S.P., and Moazed D. 2005. A Cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. *RNA Biol.* **2:** 106–111.
Horn P.J., Bastie J.N., and Peterson C.L. 2005. A Rik1-associated, cullin-dependent E3 ubiquitin ligase is essential for heterochromatin formation. *Genes Dev.* **19:** 1705–1714.
Jacob F. and Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3:** 318–356.
Jeffery L. and Nakielny S. 2004. Components of the DNA methylation system of chromatin control are RNA-binding proteins. *J. Biol. Chem.* **279:** 49479–49487.
Jia S., Kobayashi R., and Grewal S.I. 2005. Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. *Nat. Cell Biol.* **7:** 1007–1013.
Jia S., Noma K., and Grewal S.I. 2004. RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science* **304:** 1971–1976.
Kanellopoulou C., Muljo S.A., Kung A.L., Ganesan S., Drapkin R., Jenuwein T., Livingston D.M., and Rajewsky K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* **19:** 489–501.
Kanno T., Huettel B., Mette M.F., Aufsatz W., Jaligot E., Daxinger L., Kreil D.P., Matzke M., and Matzke A.J. 2005. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat. Genet.* **37:** 761–765.
Kanno T., Mette M.F., Kreil D.P., Aufsatz W., Matzke M., and Matzke A.J. 2004. Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr. Biol.* **14:** 801–805.
Kanoh J., Sadaie M., Urano T., and Ishikawa F. 2005. Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Curr. Biol.* **15:** 1808–1819.
Kato H., Goto D.B., Martienssen R.A., Urano T., Furukawa K., and Murakami Y. 2005. RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* **309:** 467–469.
Kim J.K., Gabel H.W., Kamath R.S., Tewari M., Pasquinelli A., Rual J.F., Kennedy S., Dybbs M., Bertin N., Kaplan J.M., et al. 2005. Functional genomic analysis of RNA interference in *C. elegans*. *Science* **308:** 1164–1167.
Kinoshita T., Miura A., Choi Y., Kinoshita Y., Cao X., Jacobsen S.E., Fischer R.L., and Kakutani T. 2004. One-way control of FWA imprinting in *Arabidopsis* endosperm by DNA methylation. *Science* **303:** 521–523.
Li F., Goto D.B., Zaratiegui M., Tang X., Martienssen R., and Cande W.Z. 2005. Two novel proteins, Dos1 and Dos2, interact with Rik1 to regulate heterochromatic RNA interference and histone modification. *Curr. Biol.* **15:** 1448–1457.
Li H.W. and Ding S.W. 2005. Antiviral silencing in animals. *FEBS Lett.* **579:** 5965–5973.
Lippman Z., Gendrel A.V., Black M., Vaughn M.W., Dedhia N., McCombie W.R., Lavine, K., Mittal V., May B., Kasschau K.D., et al. 2004. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430:** 471–476.
Lippman Z. and Martienssen R. 2004. The role of RNA interference in heterochromatic silencing. *Nature* **431:** 364–370.
Lippman Z., May B., Yordan C., Singer T., and Martienssen R. 2003. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol.* **1:** E67.
Maine E.M., Hauth J., Ratliff T., Vough V.E., She X., and Kelly W.G. 2005. EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during *C. elegans* meiosis. *Curr. Biol.* **15:** 1972–1978.
Martienssen R.A. 2003. Maintenance of heterochromatin by RNA interference of tandem repeats. *Nat. Genet.* **35:** 213–214.
Martienssen R.A., Zaratiegui M., and Goto D.B. 2005. RNA interference and heterochromatin in the fission yeast *Schizosaccharomyces pombe*. *Trends Genet.* **21:** 450–456.
Martienssen R., Lippman Z., May B., Ronemus M., and Vaughn M. 2004. Transposons, tandem repeats, and the silencing of imprinted genes. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 371–379.
Mathieu O. and Bender J. 2004. RNA-directed DNA methylation. *J. Cell Sci.* **117:** 4881–4888.
Matzke M., Aufsatz W., Kanno T., Daxinger L., Papp I., Mette M.F., and Matzke A.J. 2004. Genetic analysis of RNA-mediated transcriptional gene silencing. *Biochim. Biophys. Acta* **1677:** 129–141.
May B.P., Lippman Z.B., Fang Y., Spector D.L., and Martienssen R.A. 2005. Differential regulation of strand-specific transcripts from *Arabidopsis* centromeric satellite repeats. *PLoS Genet.* **1:** e79.
Mette M.F., van der Winden J., Matzke M.A., and Matzke A.J. 1999. Production of aberrant promoter transcripts contributes to methylation and silencing of unlinked homologous promoters in trans. *EMBO J.* **18:** 241–248.
Mette M.F., Aufsatz W., van der Winden J., Matzke M.A., and Matzke A.J. 2000. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* **19:** 5194–5201.
Morris K.V., Chan S.W., Jacobsen S.E., and Looney D.J. 2004. Small interfering RNA-induced transcriptional gene silencing in human cells. *Science* **305:** 1289–1292.
Motamedi M.R., Verdel A., Colmenares S.U., Gerber S.A., Gygi S.P., and Moazed D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119:** 789–802.
Napoli C., Lemieux C., and Jorgensen R. 1990. Introduction of a chimeric chalcone synthase gene into Petunia results in reversible co-suppression of homologous genes *in trans*. *Plant Cell* **2:** 279–289.
Noma K., Sugiyama T., Cam H., Verdel A., Zofall M., Jia S., Moazed D., and Grewal S.I. 2004. RITS acts in *cis* to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* **36:** 1174–1180.
Onodera Y., Haag J.R., Ream T., Nunes P.C., Pontes O., and Pikaard C.S. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120:** 613–622.
Pal-Bhadra M., Bhadra U., and Birchler J.A. 1997. Cosuppression in *Drosophila*: Gene silencing of *Alcohol dehydrogenase* by *white-Adh* transgenes is *Polycomb* dependent. *Cell* **90:** 479–490.
———. 1999. Cosuppression of nonhomologous transgenes in *Drosophila* involves mutually related endogenous sequences. *Cell* **99:** 35–46.
———. 2002. RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in *Drosophila*. *Mol. Cell.* **9:** 315–327.

Pal-Bhadra M., Leibovitch B.A., Gandhi S.G., Rao M., Bhadra U., Birchler J.A., and Elgin S.C. 2004. Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science* **303:** 669–672.

Partridge J.F., Scott K.S., Bannister A.J., Kouzarides T., and Allshire R.C. 2002. *cis*-acting DNA from fission yeast centromeres mediates histone H3 methylation and recruitment of silencing factors and cohesin to an ectopic site. *Curr. Biol.* **12:** 1652–1660.

Petrie V.J., Wuitschick J.D., Givens C.D., Kosinski A.M., and Partridge J.F. 2005. RNA interference (RNAi)-dependent and RNAi-independent association of the Chp1 chromodomain protein with distinct heterochromatic loci in fission yeast. *Mol. Cell. Biol.* **25:** 2331–2346.

Pidoux A.L. and Allshire R.C. 2004. Kinetochore and heterochromatin domains of the fission yeast centromere. *Chromosome Res.* **12:** 521–534.

Plasterk R.H. 2002. RNA silencing: The genome's immune system. *Science* **296:** 1263–1265.

Provost P., Silverstein R.A., Dishart D., Walfridsson J., Djupedal I., Kniola B., Wright A., Samuelsson B., Radmark O., and Ekwall K. 2002. Dicer is required for chromosome segregation and gene silencing in fission yeast cells. *Proc. Natl. Acad. Sci.* **99:** 16648–16653.

Ptashne M. 1967. Specific binding of the lambda phage repressor to lambda DNA. *Nature* **214:** 232–234.

Reinhart B.J. and Bartel D.P. 2002. Small RNAs correspond to centromere heterochromatic repeats. *Science* **297:** 1831.

Reiss D., Josse T., Anxolabehere D., and Ronsseray S. 2004. *aubergine* mutations in *Drosophila melanogaster* impair P cytotype determination by telomeric *P* elements inserted in heterochromatin. *Mol. Genet. Genomics.* **272:** 336–343.

Ronemus M. and Martienssen R. 2005. RNA interference: Methylation mystery. *Nature* **433:** 472–473.

Sarot E., Payen-Groschene G., Bucheton A., and Pelisson A. 2004. Evidence for a *piwi*-dependent RNA silencing of the *gypsy* endogenous retrovirus by the *Drosophila melanogaster flamenco* gene. *Genetics* **166:** 1313–1321.

Shiu P.K., Raju N.B., Zickler D., and Metzenberg R.L. 2001. Meiotic silencing by unpaired DNA. *Cell* **107:** 905–916.

Sigova A., Rhind N., and Zamore P.D. 2004. A single Argonaute protein mediates both transcriptional and posttranscriptional silencing in *Schizosaccharomyces pombe*. *Genes Dev.* **18:** 2359–2367.

Sijen T. and Plasterk R.H. 2003. Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature* **426:** 310–314.

Sijen T., Fleenor J., Simmer F., Thijssen K.L., Parrish S., Timmons L., Plasterk R.H., and Fire A. 2001. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107:** 465–476.

Soppe W.J., Jacobsen S.E., Alonso-Blanco C., Jackson J.P., Kakutani T., Koornneef M., and Peeters A.J. 2000. The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell.* **6:** 791–802.

Sugiyama T., Cam H., Verdel A., Moazed D., and Grewal S.I. 2005. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci.* **102:** 152–157.

Ting A.H., Schuebel K.E., Herman J.G., and Baylin S.B. 2005. Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. *Nat. Genet.* **37:** 906–910.

Tran R.K., Zilberman D., de Bustos C., Ditt R.F., Henikoff J.G., Lindroth A.M., Delrow J., Boyle T., Kwong S., Bryson T.D., et al. 2005. Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in *Arabidopsis*. *Genome Biol.* **6:** R90.

Vaughn M.W. and Martienssen R.A. 2005. Finding the right template: RNA Pol IV, a plant-specific RNA polymerase. *Mol. Cell* **17:** 754–756.

Verdel A., Jia S., Gerber S., Sugiyama T., Gygi S., Grewal S.I., and Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303:** 672–676.

Verdel A. and Moazed D. 2005. RNAi-directed assembly of heterochromatin in fission yeast. *FEBS Lett.* **579:** 5872–5878.

Volpe T.A., Kidner C., Hall I.M., Teng G., Grewal S.I., and Martienssen R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297:** 1833–1837.

Volpe T., Schramke V., Hamilton G.L., White S.A., Teng G., Martienssen R.A., and Allshire R.C. 2003. RNA interference is required for normal centromere function in fission yeast. *Chromosome Res.* **11:** 137–146.

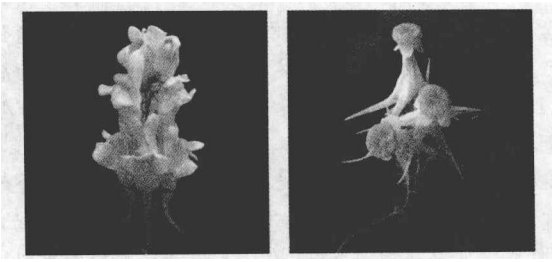
Wassenegger M., Heimes S., Riedel L., and Sanger H.L. 1994. RNA-directed de novo methylation of genomic sequences in plants. *Cell* **76:** 567–576.

Zamore P.D. 2002. Ancient pathways programmed by small RNAs. *Science* **296:** 1265–1269.

Zamore P.D., Tuschl T., Sharp P.A., and Bartel D.P. 2000. RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101:** 25–33.

Zhong W., Feng H., Santiago F.E., and Kipreos E.T. 2003. CUL-4 ubiquitin ligase maintains genome stability by restraining DNA-replication licensing. *Nature* **423:** 885–889.

# 第 9 章



# 植物的表观遗传调控

Marjori Matzke and Ortrun Mittelsten Scheid

*Gregor Mendel Institute of Molecular Plant Biology, Austrian Academy of Sciences, A-1030 Vienna, Austria*

## 本章目录

1. 植物作为表观遗传学研究对象的优点
   - 1.1 植物和哺乳动物在表观基因组结构上是相似的
   - 1.2 植物为表观遗传研究提供了新的课题
   - 1.3 哺乳动物中无法实施的研究方法，在植物中成为可能
   - 1.4 植物对表观遗传研究的贡献有目共睹
2. 植物染色质的分子组成
   - 2.1 植物中 DNA 甲基化的调控因子
   - 2.2 组蛋白修饰酶
   - 2.3 其他染色质蛋白
3. RNAi 介导的基因沉默通路的分子组成
   - 3.1 细述植物中 RNAi 介导的基因沉默
   - 3.2 通路一：转基因相关的转录后沉默及病毒诱导的基因沉默（PTGS/VIGS）
   - 3.3 通路二：通过 miRNA 和反式作用 siRNA对植物发育的调控
   - 3.4 通路三：转基因相关的转录水平沉默，RNA 指导的 DNA 甲基化和异染色质的形成
4. 没有 RNA 参与的表观遗传调控
5. 展望

参考文献

## 概要

植物是表观遗传调控的“大师”。植物中不仅有真核生物所有的表观遗传机制，还能将其发挥到其他生物无法企及的程度。通常与基因沉默相关的DNA甲基化，在植物基因组的CpG、CpNpG和CpNpN中都有发现，DNA甲基化还依赖很多植物特异性的蛋白，包括那些专门参与去甲基化过程的。组蛋白修饰酶调控植物染色质结构，其具有催化功能的区段保守性很高，但是它们的基因往往是庞大的家族，这就使得基因的功能更加广泛。植物体内RNA干扰介导的基因沉默通路功能全面，包括对抗病毒、抑制转座子、调控发育进程和加强基因组组织性的功能。尽管人们已经认识到DNA甲基化和组蛋白修饰之间存在相互作用，但近期发现这些修饰可以通过RNAi的机制作用于基因组特定的区域，这个发现为表观遗传学研究开辟了新的方向。这些沉默通路的部分重叠和相互作用给植物的表观遗传调控提供了多层次和高弹性的方式。

植物表观遗传的显著性体现在其进化史、发育模式和“生活方式”上。多倍体——染色体数目成倍的增加是植物世系中常见的现象，其作用在于扩增基因家族和保持复制基因功能上的特异性。与哺乳动物不同的是，植物的生长发育是通过分生组织（meristem）——一种个体特有的持续保持分裂功能的干细胞群，分裂产生新的地下和地上部分，而不是哺乳动物的发育模式即器官和组织的形成几乎是在胚胎发育时期完成的。因为植物这种特有的发育模式，胚胎后的发育很大程度上受环境因素的影响，因而植物发育具有很大的可塑性和多样性。由于植物无法主动脱离它所生长的环境，所以只能被迫去适应多变的甚至并不理想的生存环境。表观遗传调控机制固有的灵活性为基因活性的快速转变以及调整基因表达模式提供了便利，使得植物可以在不可预知的环境中生存和繁殖。

从科学史上看，植物为发现和分析表观遗传现象提供了绝好的研究系统。卡尔·林奈于18世纪发现的一些名为柳穿鱼（*Linaria vulgaris*）的植物（见题图）可以从镜面对称转变成辐射对称的现象后来被证实就是*cycloidea*基因——一种与花发育相关的基因表观遗传调控的结果。过去5年来在表观遗传领域的进展更是令人欣喜，拟南芥（*Arabidopsis thaliana*）因基因组测序的全部完成而成为十分重要的遗传研究工具；动物和真菌表观遗传现象的平行研究也为表观遗传学的突破提供了很好的契机。

## 1. 植物作为表观遗传学研究对象的优点

### 1.1　植物和哺乳动物在表观基因组结构上是相似的

在生物学被划分为独立的学科后不久，动物和植物就被分立。传统意义上，生物学家也根据其研究领域划分为动物学家和植物学家。当然，也有人对生物的分类方式提出了别的看法，例如异养（heterotrophic，需要外源有机物质）对自养（autotrophic，自己供能），移动生活方式对固着生活方式，潜在位移和灵活的细胞形态对固定刚性的细胞形态。然而，近几十年来，遗传学家和分子生物学家发现了动植物之间很大的相似性，虽然他们在进化上很早便分离。这些相同之处包括通过减数分裂和受精作用的有性生殖，由基因主导的个体发育调控，相关因子调控细胞分裂和增殖以及通过相似的信号级联完成对环境的适应。这种相似点已经扩展到了基因组的很多方面以及表观基因的组织。

动植物之间的相似性尤其突出地表现为二者具有相似的基因组大小、基因组复杂度以及异染色质比例的。就像其他真核生物一样，常染色质和异染色质由组蛋白特异性的乙酰化及甲基化定性，但是植物和哺乳类动物的异染色质还有另一个特点，就是大量的 DNA 胞嘧啶甲基化。通过不同的模式生物在基因组结构和表观遗传调控方面的比较，我们发现植物和哺乳动物之间的相似性甚至超过动物界本身（表 9-1）。因此，即便对课题的兴趣是以人类为中心而展开的，相似的问题也可以在植物和哺乳动物中得到回答，基本的信息在这两个系统之中都是相通的。

### 1.2　植物为表观遗传研究提供了新的课题

除了与哺乳动物共有一些基因特性外，植物本身具有独特的表观遗传现象。在哺乳动物中，受精作用是通过两个单倍体细胞融合而成，而单倍体细胞本身则是减数分裂的产物。在植物的世代交替史中，在减数分裂和受精作用之间也有单倍体（配子体，gametophyte）生长阶段（图 9-1）。植物的配子体对应的是花粉（雄性）和胚珠（雌性），每个配子体中含有数个单倍体核，它们是通过正常细胞的减数分裂形成两个单倍体细胞，而后单倍体本身再发生有丝分裂形成的。在配子体阶段，任何遗传或者表观遗传信息的丢失都是不可以通过同源染色体或等位基因来补偿的。尽管深入的研究并未展开，但是并没有证据表明植物的配子体阶段会发生大量程序性的表观遗传标记的丢失，而哺乳动物中却有这样的现象，这或许可以解释为什么在植物中表观遗传现象是可以通过减数分裂遗传给后代的。

植物另一个很突出的特点是生殖细胞系的定义并不严格以及生殖细胞系在发育晚期才与体细胞有了区别（图 9-1）。植物含有顶端分生组织，这些是茎尖和根尖的生长点，可以产生新的组织和器官。茎尖分生组织最终形成花并产生有性生殖所必须的配子，但是其他侧边的分生区也可以长出花朵。很多植物发育产生特殊的器官，如根状茎、块状茎或者鳞茎包含分生组织。这些植物无性繁殖的机制比种子发芽更常见或更成功。植物

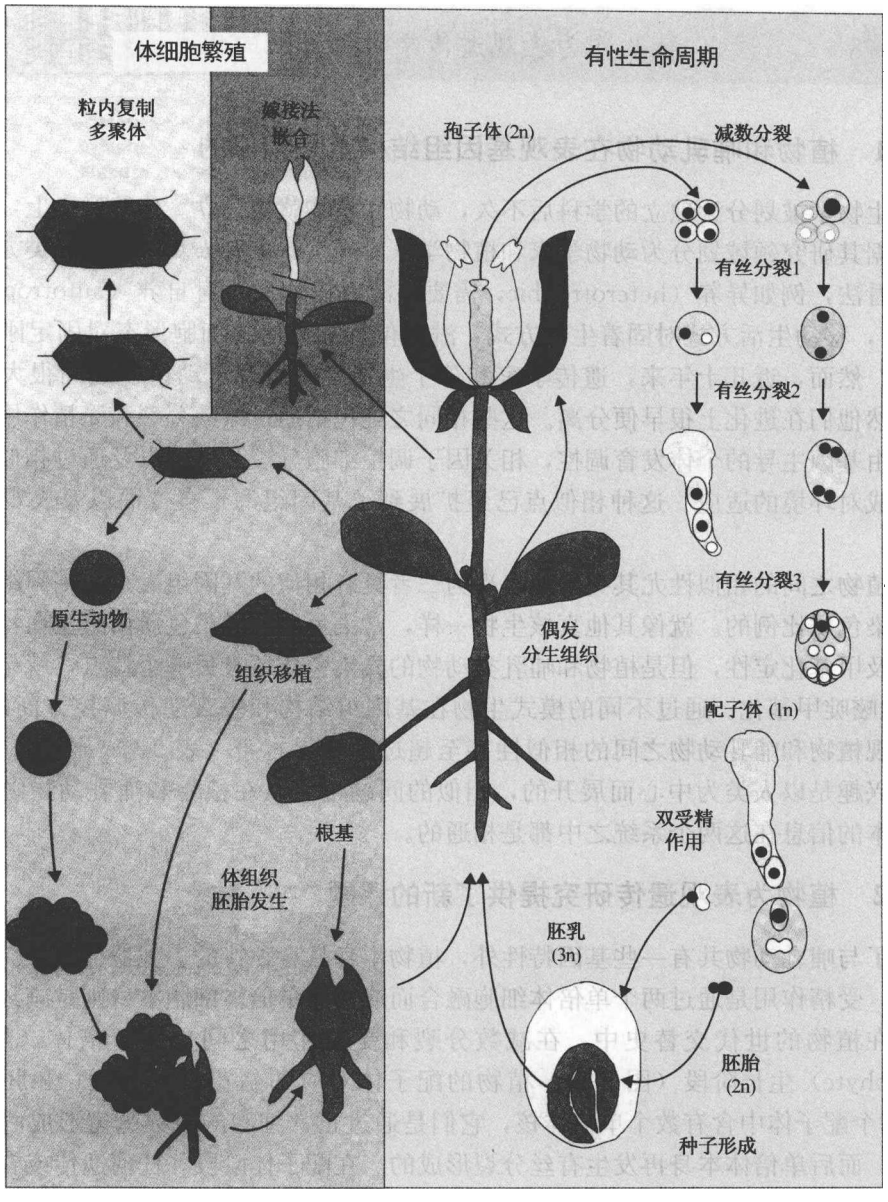



图 9-1 植物生命周期的特殊性

植物能进行有性繁殖（配子形成、受精和种子形成，右）以及体细胞繁殖（无性发芽，去-和再分化或胚芽形成，左）。高等植物有根、茎、叶和花，是二倍孢子体。减数分裂期间染色体数目减少到一半。尽管动物减数分裂产物形成配子无需继续分裂而直接融合产生二倍体胚胎，植物则依靠 2 或 3 次有丝分裂分别形成单倍体雄性或雌性配子体。花粉管最终含有 1 个营养核（白色）和 2 个传代核（黑色）。这两个传代核使卵细胞（黑色）和中心细胞受精，中心细胞有着由两个极核（黄色）融合而成的二倍体核。这两次受精产生双倍体的胚芽和三倍体胚乳，胚乳则为胚芽发育提供营养来源。种子萌芽后胚芽将长成为新的孢子体。另外，多数植物有无性繁殖的潜能可以通过诸如激活静止的侧向分裂、根的结构特化成管状、组织培养的扩增、甚至个别体细胞在细胞壁移除后的原生质再生等方式进行。植物中内源性复制是很频繁的，能产生多倍体细胞或组织。植物能被嫁接产生嵌合体。总之，植物的遗传和表观遗传信息因此经过了比动物生殖细胞系更粗的途径。

胚胎不仅仅由受精卵发育而来，也可以从体细胞组织发育而来（称之为体细胞胚胎化）。通过对植物组织的培养操作，一些分化了的体细胞可以去分化重新进入分化阶段。这意味着在哺乳动物中成功率极低的体细胞克隆，却是很多植物的常规操作，也表明这些年里已经产生了数不尽的“绿色多莉”。尽管如此，在这些理论上说遗传组成上完全相同的植物群落中却惊奇地发现了表型不同的植株。这种现象被称为“无性系变异”（somaclonal variation），无性系变异具有很强的表观遗传学基础，对于植物的生长和环境适应很有帮助（综述见 Kaepper et al. 2000）。

植物独具的另一个特点是胞间连丝，也就是个体细胞胞质间的桥梁，胞间连丝对小分子，一些蛋白、RNA 及病毒基因组都具有通透性。除了这种高度的细胞间作用，植物的茎可以嫁接到另一株遗传组成上并不相同的植物的顶端并长出幼芽（图 9-1）。这种技术产生了嫁接植物，也就是植物本体的组织与产生花粉、胚囊的组织在遗传上是不同的。因此，尽管表观遗传标记是通过生殖细胞传递到后代的，在植物中这种传递相比动物而言更灵活了。

**表 9-1　表观遗传学模式生物的基因组和表观遗传组成列表**

| 特点 | 酿酒酵母 | 粟酒裂殖酵母 | 链孢霉 | 线虫 | 果蝇 | 哺乳动物 | 植物 |
|---|---|---|---|---|---|---|---|
| 基因组大小 | 12Mb | 14Mb | 40Mb | 100Mb | 180Mb | 3 400Mb | 150～5 000Mb |
| 基因数量 | 6 000 | 5 000 | 10 000 | 20 000 | 14 000 | 25 000～30 000 | 25 000～40 000 |
| 基因的平均大小 | 1.45kb | 1.45kb | 1.7kb | 2kb | 5kb | 35～46kb | 2kb |
| 内含子/基因的平均数量 | ≤1(4%基因有内含子) | 2(40%基因有内含子) | 2 | 5 | 3 | 6～8 | 4～5 |
| 编码蛋白的基因组比例 | 70 | 60 | 44 | 25 | 13 | 4(Hs) | 26(At)<br>10(Os) |
| 转座子沉默 | (+) | + | +(+RIP)[a] | + | + | + | + |
| 基因印记 | — | — | — | — | +[b] | + | + |
| RNA 干扰机制 | — | + | + | + | + | + | + |
| 抑制型组蛋白甲基化 | — | H3K9(+)<br>— | H3K9+<br>H3K27+ | H3K9+<br>H3K27+ | H3K9+<br>H3K27+ | H3K9+<br>H3K27+ | H3K9+<br>H3K27+ |
| DNA 甲基化(CG 上) | — | — | + | — | — | + | + |
| DNA 甲基化(CNG/CNN 上) | — | — | + | — | (+) | (−) | + |
| 与 DNA 甲基化及识别相关的基因 | — | +[c] | + | +[d] | +[e] | + | + |
| HP1 相似蛋白 | — | + | + | + | + | + | + |
| PcG 蛋白 | — | — | — | + | + | + | + |

注：(At) 拟南芥；(Cb) 一种线虫；(Ce) 秀丽隐杆线虫；(Dm) 果蝇；(Hs) 人；(Os) 水稻；(Pp) 圆型线虫

a 重复诱导的点突变，详见第 6 章

b 染色体或者基因组范围，而不是特定基因而言

c 突变的 Dnmt2

d Dnmt2(Pp) 和 MBD 结构域蛋白 (Ce、Cb、Pp)

e Dnmt2 和 MBD 结构域蛋白 (Dm)

植物比哺乳动物对多倍性（染色体的加倍）的耐受更高。野生的大量多倍体物种以及人工培育的多倍体植物，如小麦、棉花、土豆、花生、甘蔗和烟草表明多倍体植物在竞争力方面具有优势。通过对植物基因组序列，包括较小的拟南芥基因序列的分析，人们可以推断出其原始序列以及在进化过程中的基因复制事件。甚至二倍体的植物也可以含有多倍体的细胞，然而这种现象在哺乳动物中十分罕见。多倍体的形成往往与基因组的大幅改变和表观遗传的改变密切相关（综述见 Adams and Wendel 2005）。其中的一些变化发生在同一代或者几代之内，对于植物的快速适应环境和进化都有很大的贡献。

## 1.3 哺乳动物中无法实施的研究方法，在植物中成为可能

在哺乳动物中，遗传学方法局限于构造各种突变体并不断的杂交产生纯合体，因为只有纯合体才能显出隐性表型。相反的，植物构建突变体的方法十分有效，通过化学、物理手段或者大量随机插入带标签的转基因或转座子都可以完成。纯合突变的个体可以通过简单的自交法，并按照孟德尔定律获得相应比例的纯合个体。筛选表观遗传调控因子主要是基于一些标记基因的激活表达或者是一些已激活报告基因的表达下调。除了这种没有争议的前导性方法，特定基因组整合位点的插入突变体迅速增加，使人们可以通过分析与其他模式生物表观调控子同源的特定基因突变效应进行逆向遗传研究（表 9-2）。

**表 9-2 通过正向和反向基因筛选途径发现的模式植物拟南芥的表观遗传调控因子组成**

| 基因或突变体(缩写) | 基因或突变体名称 | 蛋白类型、确认或预测的功能 | 突变体筛选 |
|---|---|---|---|
| DNA 甲基化 | | | |
| *MET1* | methyltransferase | DNA 甲基转移酶(CG) | 内源重复基因的再激活 |
| *DDM2* | decreased DNA methylation | | 转基因的再激活<br>RdDM 的干扰<br>中心粒区域的低甲基化 |
| *CMT3* | chromomethyltransferase | DNA 甲基转移酶(非 CG) | SUP-clk 的再激活<br>PAI 的再激活 |
| *DRM1* | domain-rearranged methyltransferase | DNA 重新甲基转移酶 | 寻找序列插入突变体 |
| *DRM2* | domain-rearranged methyltransferase | DNA 重新甲基转移酶 | 寻找序列插入突变体 |
| *HOG1* | homology-dependent gene silencing | S-腺苷同型半胱氨酸水解酶 | 重复性转基因的再激活 |
| *ROS1* | repressor of silencing | DNA 糖基酶结构域蛋白 | 转基因的失活 |
| *DME* | demeter | DNA 糖基酶结构域蛋白 | 败育 |
| 组蛋白修饰 | | | |
| *HDA1* | histone deacetylase | 组蛋白去乙酰化酶 | 寻找序列插入突变体<br>反义表达 |
| *HDA6* | histone deacetylase | 组蛋白去乙酰化酶 | 重复性转基因的激活 |
| *SIL1* | modifier of silencing | | |
| *AXE1* | auxin-gene repression | | 干扰 RdDM |
| *RTS1* | RNA-mediated transcriptional silencing | | |

续表

| 基因或突变体(缩写) | 基因或突变体名称 | 蛋白类型、确认或预测的功能 | 突变体筛选 |
|---|---|---|---|
| 组蛋白修饰 | | | |
| *SUVH2* | Su(var)3-9 homolog | 组蛋白甲基化酶 | 寻找序列插入突变体<br>反义表达 |
| *SUVH4* | Su(var)3-9 homolog | 组蛋白甲基化酶 | PAI 再激活 |
| *KYP1* | kryptonite | | SUP-clk 的再激活 |
| 染色质形成/重塑 | | | |
| *DDM1* | decreased DNA methylation | SWI2/SNF2 ATP 酶 | 中心粒重复区域的低甲基化 |
| *SOM* | Somniferous | | 重复性转基因的激活 |
| *DRD1* | defective in RNA-directed DNA methylation | SWI2/SNF2 ATP 酶 | 用 RdDM 干扰 |
| *SPD* | splayed | SWI2/SNF2 ATP 酶 | 改变的分生组织的维持 |
| *PIE* | photoperiod-independent early flowering | 依赖 ATP 的染色质重塑蛋白 | 开花时间的变化 |
| *PKL* | pickle | CHD3 染色质重塑因子 | 异常的根部发育 |
| *FAS1* | fasciated | 染色质组装因子亚基 | 形态变化 |
| *FAS2* | fasciated | 染色质组装因子亚基 | 形态变化 |
| *BRU1* | brushy | 未标明特性的蛋白 | DNA 损伤的敏感性 |
| *LHP1* | like heterochromatin protein | 抑制性染色质的形成 | 早开花及形态改变 |
| *MOM1* | morpheus' molecule | 不完全的 SWI2/SNF2 ATP 酶 | 重复性转基因的再激活 |
| *RPA2* | replication protein A | ssDNA 的复制蛋白复合体亚基 | Ros1 抑制筛选<br>重复性转基因的再激活 |
| RNA 干扰介导的基因沉默 | | | |
| *DCL1* | Dicer-like | RNA 酶 III(双链 RNA 酶) | 寻找插入序列突变体 |
| *CAF1* | carpel factory | | 不正常的花发育 |
| *SIN1* | short integuments | | 不正常胚珠发育 |
| *EMB76* | embryo-defective | | 胚胎发育不良 |
| *SUS1* | suspensor | | 胚柄增殖 |
| *DCL2* | Dicer-like | RNA 酶 III(双链 RNA 酶) | 寻找插入序列突变体 |
| *DCL3* | Dicer-like | RNA 酶 III(双链 RNA 酶) | 寻找插入序列突变体 |
| *AGO1* | argonaute | PAZ-PIWI 结构域蛋白 | 形态改变<br>转基因再激活 |
| *AGO4* | argonaute | PAZ-PIWI 结构域蛋白 | SUP-clk 再激活 |
| *AGO7* | argonaute | PAZ-PIWI 结构域蛋白 | 表皮毛发育的时间和类型 |
| *ZIP* | zippy | | |

续表

| 基因或突变体(缩写) | 基因或突变体名称 | 蛋白类型、确认或预测的功能 | 突变体筛选 |
|---|---|---|---|
| RNA 干扰介导的基因沉默 | | | |
| *AGO10* | argonaute | PAZ-PIWI 结构域蛋白 | 分生组织缺陷 |
| *PNH/ZLL* | | | |
| *RDR1* | RNA-dependent RNA polymerase | RNA 依赖的 RNA 聚合酶 | 寻找插入序列突变体 |
| *RDR2* | RNA-dependent RNA polymerase | RNA 依赖的 RNA 聚合酶 | 寻找插入序列突变体 |
| *RDR6* | RNA-dependent RNA polymerase | RNA 依赖的 RNA 聚合酶 | 转基因再激活<br>毛状体发育改变 |
| *SDE1* | silencing-defective | | |
| *SGS2* | suppressor of gene silencing | | |
| *NRPD1a* | RNA polymerase | RNA 聚合酶 IV 亚基 | 转基因再激活 |
| *SDE4* | silencing-defective | | |
| *NRPD1b* | RNA polymerase | RNA 聚合酶 IV 亚基 | 用 RdDM 干扰 |
| *DRD3* | defective in RNA-directed DNA methylation | | |
| *NRPD2a* | RNA polymerase | RNA 聚合酶 IV 亚基 | 寻找插入序列突变体 |
| *DRD2* | defective in RNA-directed DNA methylation | | 用 RdDM 干扰 |
| *SDE3* | silencing-defective | RNA 解旋酶 | 转基因再激活 |
| *SGS3* | suppressor of gene silencing | 超螺旋蛋白 | 转基因再激活<br>毛状体发育改变 |
| *HEN1* | HUA enhancer | 结合双链 RNA,甲基转移酶 | 转基因再激活<br>形态改变 |
| *HYL1* | hyponastic leaves | 双链 RNA 结合的核蛋白 | 形态改变<br>寻找插入序列突变体 |
| *WEX* | Werner syndrome-like exonuclease | RNA 酶 D 核酸外切酶 | 寻找插入序列突变体 |
| *XRN4* | XRN homolog | 核糖核酸外切酶 | 寻找插入序列突变体 |
| *HST* | hasty | miRNA 输出受体 | 毛状体发育的时间和类型 |
| Polycomb 家族 | | | |
| *MEA* | medea | Polycomb 家族 | 败育 |
| *FIS1* | fertilization-independent seeds | | 无需受精的种子形成 |
| *CLF* | Curly leaf | Polycomb 家族 | 叶子形态改变 |
| *FIE* | fertilization-independent endosperm | Polycomb 家族 | 无需受精的种子形成 |
| *FIS3* | fertilization-independent seeds | | |
| *MSI1* | multicopy suppressor of IRA homolog | Polycomb 家族<br>染色质组装因子亚基 | 寻找插入序列突变体<br>胚乳模式缺陷 |
| *SWN* | swinger | Polycomb 家族 | 寻找插入序列突变体 |
| *EMF2* | embryonic flower | Polycomb 家族 | 缺少营养阶段 |
| *VRN2* | vernalization | Polycomb 家族 | 春化却晚开花 |

同源基因家族成员的数量在植物和哺乳动物中有很大的差异。因为功能性重复基因的存在，有些突变体对植物或哺乳动物的影响不再那么严重，这种现象对研究而言非常重要，因为有些基因的功能完全丧失功能是会使相应个体致死而无法在早期发育阶段做分析，而如果其影响不那么大，则可能使其发育功能研究成为可能。在体内表达决定植物组织颜色的信号通路中，并非必需的基因是非常容易的（图 9-2a～e）。其他表观遗传变化可以引起形态缺陷，这种形态缺陷不会引起致命的后果（图 9-2f）。另外，在亲核内源或转基因性状表达的选择实验中，又可筛选出成千上万的个体植株（图 9-2g）。

拟南芥因全基因序列已知，大型插入突变体库已建立以及简便的突变体操作手段而成为目前表观遗传学研究的模式生物。接下来的内容将讲述如何通过转基因沉默的手段从突变体中筛选表观遗传调控子，这些方法包括使用生物信息学的方法识别其他生物沉默基因的同源基因，也包括用正向或反向的遗传方法筛选其他的因沉默导致的缺陷的突变体中相似的发育表型（表 9-2）。如果需要得到拟南芥、水稻和玉米染色体基因更全面详细的信息，读者可以登录植物染色质数据库（Plant Chromatin Database，http://www.chromdb.org）查询。

## 1.4 植物对表观遗传研究的贡献有目共睹

证明植物模式系统在表观遗传学研究中价值的最好证据就是植物学家们在这个领域所做的大量杰出的奠基性贡献。常染色质和异染色质的区别首先是在苔藓细胞学研究中发现的（Heitz 1928）。在土豆和玉米中发现等位基因相互作用后发生的可遗传但可回复的基因表达变化，这种现象之后被命名为副突变（paramutation），这也是非孟德尔遗传定律的早期证据（综述见 Chandler and Stam 2004），这种现象后来也在哺乳动物中发现。同样的，个体基因的亲本印记现象（parental imprinting）是首先在玉米中发现的（综述见 Alleman and Doctor 2000）。个体反复发生有对称性改变的花，被卡尔·林奈称为“反常整齐花”（peloria）（见题图），现在可以用表观等位基因（epiallele）的形成来解释，表观等位基因是一种基因稳定的表观遗传修饰，而基因序列本身并没有发生变化（Cubas et al. 1999）。植物细胞学研究表明次级染色体浓缩的变化与核仁显性相关并依赖表观遗传调控，核仁显性现象发生在物种间杂合体上，表现为一方亲本的 rRNA 沉默（综述见 Pikaard 2000）。玉米转座子的开拓性研究是由 Barbara McClintock 和其他一些研究人员完成的，该项研究表明玉米的遗传行为与其表观遗传调控有着紧密的联系（综述见 Fedoroff and Chandler 1994）。事实上，植物中现存的转座子和早期转座子的退化形态仍旧为植物基因组上的表观遗传调控提供了基础（本章 3.4 节）。

当 20 世纪 80 年代末期，对植物例如烟草、矮牵牛花和拟南芥的转基因技术成为常规以后，表观遗传研究因为标记基因的引入而获得更进一步的发展（综述见 Jorgensen 2003；Matzke and Matzke 2004）。由于基因沉默越发明显地与多拷贝连锁或非连锁转基因密切相关，同源性依赖的基因沉默的概念渐渐形成。不同的同源性依赖基因沉默主要由 2 个原因导致：mRNA 关闭（转录后基因沉默，posttranscriptional gene silencing，PTGS）和转录抑制（转录水平基因沉默，transcriptional gene silencing，TGS），这两者都与沉默基因的胞嘧啶的甲基化相关。转基因矮牵牛花的转录后基因沉默有一个著名

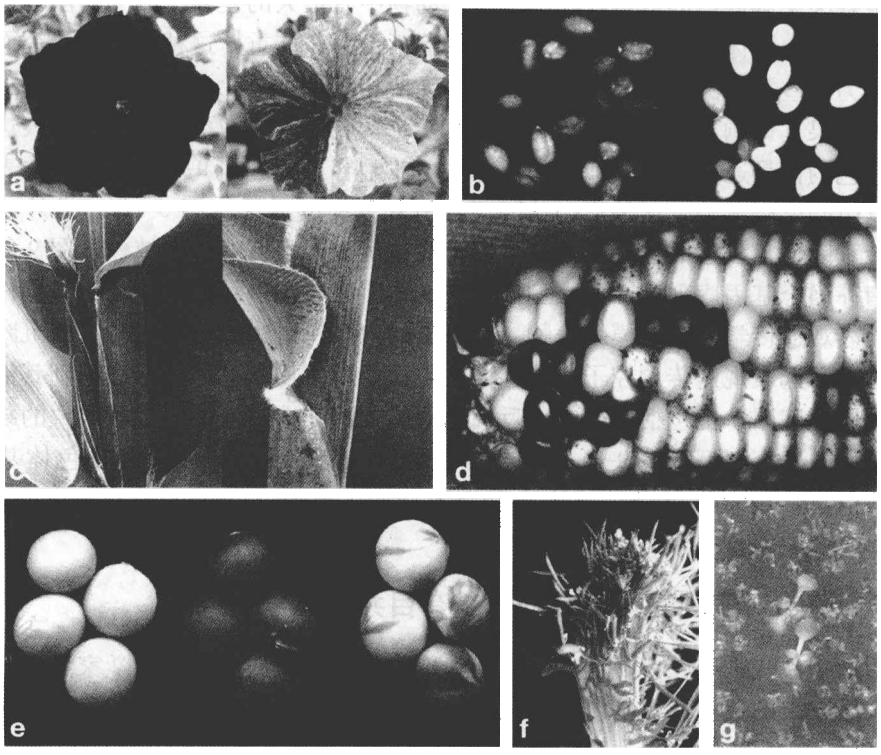



图 9-2 植物的表观遗传控制实验分析

决定植物组织颜色的基因可进行简易并低成本的体内表达实验。(a) 牵牛花的深紫色需要黄烷酮醇还原酶（*dihydroflavonol reductase*，*DFR*）基因表达，而 *DFR* 基因沉默会产生杂而浅亮的颜色（Jan Kooter）。(b) 拟南芥种子表达有苯丙烯酰苯合成酶（*Chalcone synthase*，*CHS*），呈深色的种子外皮，但是 *CHS* 在整合了同源转基因而基因沉默后呈黄色种子（Lan Furner 经许可转自）。(c) 有 *B-I* 基因表达的玉米植物含有紫色色素，区别于绿色植物中副突变的、失活的 *B′*基因。(d) 玉米穗，*B-Peru* 基因中分离转座子插入（segregating a transposon insertion，*Spm*）元件为花青苷色素所需。紫色玉米粒代表 1 种回复突变体，其 *Spm* 元件从生殖细胞的 *B-Peru* 基因中被切除。有黑点的玉米粒含有活性 *Spm* 元件，其在玉米粒的发育期间不断地诱导体细胞的部分切除。玉米粒有较少的，小的紫色的部分说明玉米粒的 *Spm* 元件被表观遗传沉默了（经许可转自 Vicky Chandler）。(e) 暗色大豆（中）是培养的变体（左）由于 *CHS* 基因自然的转录后翻译沉默而抑制色素所致，它可由亲本植物通过携有 PTGS 抑制蛋白的病毒转染而部分得以回复并呈杂色模式（右）。(f) 表观遗传调控在植物形态学功能上也是显然的：功能削弱的染色质组装因子亚单位可导致簇生的拟南芥茎。(g) 拟南芥转基因标记的沉默释放可通过在选择性培养基中的生长来计数。(c，经许可转自 Chandler et al. 2000 [©Springer]；e，经许可转自 Senda et al. 2004 [©美国植物生物学家学会])。

的例子，最初被称为“共抑制”，这里有一个具体的例子：原本研究者试图通过过表达查尔酮合成酶基因（*chalcone*，*CHS*）来改变花朵的颜色，过表达该基因后会使花瓣呈现紫色，然而实验结果却出现杂色甚至是完全白色的花朵。这种色素的缺失现在被证实是因为 *CHS* 转基因以及 *CHS* 基因内源基因的共同失活（Jorgensen 2003）。转录后基因沉默在植物中已经被认为是和线虫 RNA 干扰等价的调控手段（本章 3.2 节）。

20 世纪 90 年代中期，转录后基因沉默与病毒抗性的关系得到进一步阐明。转录后

基因沉默能天然地保护了植物使外来的病毒不能随意复制，这些病毒就是转录后基因沉默的诱导物或靶标。该原理拓展用于实验中下调表达植物基因，通过构建包含与目标片段同源性较高的病毒载体，可以实现病毒诱导的基因沉默（virus-induced gene silencing，VIGS；综述见 Burch-Smith et al. 2004）。另外，RNA 指导的 DNA 甲基化现象也发现于病毒感染的植物中，这是 RNA 通过诱导表观遗传修饰作用于 DNA 的最好证据（Wassenegger et al. 1994）。通过目的性构建同源性的双链 RNA，该原理被成功地应用于转录水平沉默基因和启动子的甲基化（见本章第 3.4 节）。

## 2. 植物染色质的分子组成

大量的植物表观遗传调控的分子组成都是通过上述提到的拟南芥的突变方法所得到（表 9-2）。然而，突变体筛选可能无法完整地揭示所有的表观遗传修饰，这是因为较大的基因家族里有功能冗余现象而重要基因的缺失可能致死。

### 2.1　植物中 DNA 甲基化的调控因子

无论是在植物还是在哺乳动物中 DNA 胞嘧啶 $C^5$ 的甲基化都是表观遗传失活或异染色质的标志。（表 9-1）（另见第 18 章）。然而在植物中，DNA 甲基化在其模式、机制蛋白上都有独特之处，而且还在非分裂的细胞中去除甲基化（综述见 Chen et al. 2005）。在本节中，我们将着重阐述与产生、维持、转译和去甲基化相关的蛋白。RNA 指导的 DNA 甲基化过程所需要的特殊分子组件将在本章 3.4 节中阐述。

#### DNA 甲基转移酶

DNA 甲基化可以被分为两步：从头甲基化和维持甲基化。从头甲基化指的是对非甲基化的 DNA 序列进行的修饰（图 9-1）（第 18 章）。在植物中从头甲基化的部位为 CpG、CpNpG 和 CpNpN 核酸组（此处 N 代表 A、T 或 C）。相反的，哺乳动物中的甲基化大多局限于 CpG 二核苷酸序列，对于不对称的 CpNpN 核酸组合没有证据显示可以甲基化。尽管导致从头甲基化的信号大多未知，但在植物中双链 RNA 却可以实现这个功能（本章 3.4 节）。维持甲基化是在 DNA 复制期间维持 DNA 甲基化的状态，通常发生在 CpG 和 CpNpG 这样回文对称的核酸序列上。甲基化的维持发生在复制或修复后半甲基化的 DNA 底物上，亲本链依旧存在的修饰可以指导维持甲基化的发生。尽管有假设认为不同的 DNA 胞嘧啶甲基转移酶完成不同类型的甲基化过程，但是植物研究中出现一种新的看法：即特异作用位点的酶（CpG 或者非 CpG）合作同时催化两个步骤。

DNA 甲基转移酶的 3 个保守家族都存在于植物中。甲基转移酶 I（methyl transferase 1，MET1）家族的成员与哺乳动物中 Dnmt1 的基因有同源性（见第 18 章），被认为参与 CpG 位点的新甲基化，然而，其家族中有一成员也参与了 CpG 位点由 RdDM 通路指导的从头甲基化的过程（本章 3.4 节）。Dnmt2 类甲基转移酶中的成员之一为拟南芥的基因组所编码，Dnmt2 类甲基转移酶几乎包含所有普遍的、保守性很高的 DNA 甲基化转移酶家族（表 9-2），然而其功能一直无法得到清晰揭示。植物的结构域重排甲

基转移酶（domains-rearranged methyltransferase，DRM）和其哺乳动物同源家族 Dnmt3 通常被认为是从头甲基化的甲基转移酶。DRM 家族的酶催化整个序列的胞嘧啶甲基化，是 RdDM 通路中非常重要的组成成分（本章 4.4 节）。就如同其名字讲述的那样，DRM 家族蛋白可以重排结构域（后面接着 I-V 的 VI-X 构象），而 Dnmt3 家族是 I-X 构象的。正因如此，这个家族的酶才有特殊的能力去甲基化像 CpNpN 这样不对称的核酸序列，而哺乳动物中这样的序列是无法甲基化的。植物特有的染色体甲基化酶 CMT3 可以修饰 CpNpG 三核苷酸。与 MET1 相似的是，CMT3 同时参与从头甲基化和维持甲基化的过程。CMT3 的准确功能尚不清楚，尽管 CMT3 失去功能的突变体会重新激活特定原本沉默的转座子（综述见 Chan et al. 2005）。

哺乳动物的 *dnmt1* 和 *dnmt3* 突变体在胚胎发育阶段或是生后不久会死亡。与其不同的是，植物的 *met1*、*cmt3* 和 *drm* 的突变体可以存活且能繁殖。植物在缺失 DNA 甲基转移酶时的存活能力使得科学家可以在甲基转移酶在发育和有性生殖的功能上开展更广泛的研究，其成功率要超过哺乳动物（综述见 Chan et al. 2005）。

### 主动的 CpG 去甲基化和 DNA 的糖基化

表观遗传调控表示一些激活或失活的基因状态而导致的标记是能回复的。DNA 甲基化就允许了这种回复性，因为甲基化本身可以通过主动或者被动的方式去除。被动地去甲基化发生在多轮 DNA 复制时甲基化无法继续维持的情况下。相反的，主动地去甲基化可以发生在未分化的细胞或者需要酶活性的时候。动物系统的早期研究报告表明主动地去甲基化时由 DNA 糖基化酶引发的，而 DNA 糖苷酶常参与 DNA 的碱基切除修复模式（综述见 Kress et al. 2001）。拟南芥中 Demeter（DME）和沉默抑制子（repressor of silencing，ROS1）的发现重新点燃了科学家对这个领域的兴趣，DME 和 ROS1 是含有 DNA 糖苷酶结构域的大型蛋白。ROS1 的基因是在筛选表观遗传下调和稳定表达的报告基因过甲基化时发现的（Gong et al. 2002）。ROS1 蛋白具有 DNA 缺口活性，但仅限于甲基化的 DNA 而非非甲基化的 DNA，这也与它在 DNA 碱基切除修复中移除甲基化的胞嘧啶的功能是一致的。ROS1 持续表达，因而可能与非分裂细胞中发育全过程没有 DNA 甲基化的现象有重要的联系（Kapoor et al. 2005a）。相反，DME 的活性仅限于雌性的配子体，并可以激活印记因子 *Medea*（*MEA*），该激活方式是功能性 DNA 糖苷酶结构域依赖性的（Choi et al. 2002）。CG 的甲基转移酶 MET1 与 DME 的作用相互拮抗，表明 DME 确实是 CG 二核苷酸去甲基化必须的因子（Hsieh and Fisher 2005）。在拟南芥中，还有两个非典型的 DME/ROS1 家族成员，这两个成员是植物特有的。该基因家族的扩大表明通过去甲基化达到可回复性的基因沉默是植物生理、发育和适应力上重要的手段。

### 甲基化 DNA 结合蛋白

甲基化的 CG 结合域蛋白（methyl-CG-binding domain，MBD）被认为是 DNA 甲基化模式信息转换成不同基因转录活性的途径。在哺乳动物中，MBD 蛋白与甲基化的 DNA 结合并行使多种功能促进转录水平沉默，如募集组蛋白去乙酰化酶等。拟南芥有

12个包含MBD的基因，哺乳动物中有11个，果蝇中5个，线虫中2个，而至现今测序完成的真菌中没有发现（Hung and Shen 2003）。拟南芥的MBD蛋白的功能现在知之甚少，但是如果将其中一个*AtMBD11*通过RNA干扰的方式敲除，会对发育过程产生多种影响（Springer and Kaeppler 2005）。通过正向遗传筛选的方法没有得到任何拟南芥MBD蛋白，可能因为功能冗余性的存在。另外，尽管哺乳动物和动物中DNA甲基转移酶的氨基酸保守性很高，但是包含MBD的蛋白除了甲基化的CG结合结构域外几乎没有相似的地方。因此，尽管植物和哺乳动物用相关的酶建立和维持DNA甲基，但是他们却是通过不同的方式将甲基化模式转译成信号的，不同的MBD蛋白就是很好的例证（Springer and Kaeppler 2005）。

甲基供体的合成体组成

甲基化酶需要活化的甲基基团，通常是S-腺苷甲硫氨酸的形式。让人惊奇的，提供这个辅助因子的生化途径本身与表观遗传调控没有什么联系。直到最近，才有拟南芥中的一个编码S-腺苷高半胱氨酸水解酶的基因*hog1*的突变被发现与表观遗传缺陷有关系（Rocha et al. 2005）。

## 2.2 组蛋白修饰酶

就像其他生物一样（表9-1），植物也含有转录后修饰组蛋白氨基末端的酶，因而建立了公认的组蛋白密码（综述见Loidl 2004）。在植物中，组蛋白修饰酶通常由相对较大的基因家族编码。大多家族成员的功能信息的了解还是很有限。两种最常见的修饰的形式是组蛋白乙酰化/去乙酰化和组蛋白甲基化。

组蛋白去乙酰化酶和乙酰转移酶

乙酰转移酶（histone acetyltransferases，HAT）和去乙酰化酶（histone deacetylases，HDAC）的相反作用确保了乙酰化这种表观遗传标记是可逆的。组蛋白低乙酰化和CpG甲基化频繁共存于某个沉默基因的现象更推动了这种可逆性，CpG甲基化可以被DNA糖苷酶主动除去（本章2.1节，主动的CG去甲基化和DNA糖基化）。拟南芥有18个被公认的HDAC和12个被公认的HAT（Pandey et al. 2002），这与哺乳动物体内发现的个数相同，却比其他非植物的真核生物数量多。公认的HDAC在所有的真核生物中是保守的，但是有一个植物特有的家族*HD2*，*HD2*的功能未知。遗传筛选法只发现了2个家族成员：*HDA1*和*HDA6*（表9-2）。从*HDA6*缺陷突变体的常规表型来看，*HDA6*对发育的作用很小，但是对维持由RNA诱导的DNA甲基化以及重复序列的甲基化的方面有重要的作用。相反，*HDA1*的低表达会导致发育多元化的影响。在拟南芥的正向筛选过程中，没有发现HAT家族的成员，这可能是功能冗余作用，也可能是因为多数的筛选体系使沉默基因被重新激活。

组蛋白甲基转移酶

能够甲基化组蛋白中赖氨酸残基的蛋白（参阅书中组蛋白赖氨酸甲基转移酶）和其

他蛋白具有常见的 SET 结构域（SU(VAR)/E(Z)/TRX)。由于他们具有甲基化 H3 和 H4 组蛋白上很多赖氨酸残基的能力，不同含有 SET 结构域的蛋白都具有增强或抑制特定基因的转录或是形成异染色质方面的功能。有些 SET 结构域蛋白是 PcG(Polycomb group）或 TrxG(Trithorax group）蛋白家族的成员，这些蛋白在植物和动物的发育阶段维持一些同源异型基因的转录抑制或激活状态（见第 11、12 章)。其他的 SET 结构域蛋白，例如 SU(VAR)3-9 通过 H3K9 甲基化参与维持异染色质区域的致密状态。

拟南芥的基因组编码 32 个 SET 结构域蛋白，其中 30 个是表达的。他们主要可以被分成 4 个家族：E(Z)、TRX、ASH1 以及 SU(VAR)3-9，还有只存在于酵母和植物中的很小的第五家族（Baumbusch et al. 2001；Springer et al. 2003)。拟南芥中表达的 SET 结构域蛋白的数量相比其他生物是较高的，果蝇中只表达 14 个，而裂殖酵母中只表达 4 个，老鼠更高一些，有 50 个。另外拟南芥通过多倍化，反移位等拓展 SET 结构域蛋白家族也使得 SU(VAR)3-9 成员得以扩增。除 SET 结构域外，植物及动物蛋白并不是十分保守的。结构域的多样性预示着调节蛋白-蛋白间的多种反应性。同时植物中的 SET 结构域蛋白被认为是以复合物的形式发挥作用而区别于动物中的 SET 蛋白。

尽管并不全面，有关已知的拟南芥 SET 结构域蛋白的功能信息仍提示了其在染色质调控和表观遗传方面的作用。最初通过遗传学方法筛选识别的 2 个 SET 结构域蛋白是 Curly leaf(CLF）和 Medea/ Fertilization independent seed formation(MEA/FIS1）蛋白，两者是果蝇 E/Z 蛋白的负调控子。除了作为 SET 结构域蛋白，MEA、CLF 和 E/Z 也属于 PcG 蛋白（第 2.3 节，其他 Polycomb 蛋白)。*CLF* 突变体可导致叶子形状的改变及花发育中的同源异型变化。*MEA/FIS1* 调控着配子体特异的基因表达并且作为印记因子抑制未受精状态下的内胚层发育（综述见 Schubert et al. 2005)。相反，TRX 蛋白家族成员拟南芥 Trithorax1 蛋白（ATX1）作为花的同源异型基因的激活因子，被假定有能力催化组蛋白 3 在赖氨酸 4 号位（H3K4）的甲基化，而 H3K4 通常是转录活性的染色质的标记（综述见 Hsieh and Fischer 2005)。

Kryptonite/ Suppressor of variegation 3-9 homolog 4(*KYP/SUVH4*）蛋白是在筛选两个内源基因表观遗传沉默抑制子的实验中被识别的（Jackson et al. 2002；Malagnac et al. 2002)。KYP/SUVH4 催化 H3 在 9 位赖氨酸的单、双甲基化并且与 CMT 协同作用维持拟南芥的部分序列的 CpNpG 的甲基化。KYP/SUVH4 似乎在异染色质形成中只起到微弱的作用（Chan et al. 2005）相反，识别于沉默转基因再激活筛选实验的 Suppressor of Variegation 3-9 homolog 2(SUVH2）蛋白似乎以激活拟南芥异染色质的 H3K9 和 H3K27 的甲基化为主（Naumann et al. 2005)。

组蛋白 H3 和 H4 的赖氨酸可被单、双或三甲基化，进而提高甲基化修饰的组合复杂性。不同生物体中特定的甲基化状态可定义异染色质。例如，H3K9me3 是动物和真菌的异染色质的主要特征，而在拟南芥则是常染色质的表观遗传标志。相反的，H3K9me1 和 H3K9me2 是拟南芥异染色质沉默的主要标志，而在哺乳动物却是常染色质修饰的标志。这些差异的缘由和他们如何与所涉及的假定的组蛋白密码相关联的机制有待探讨。另外，植物中特异的组蛋白修饰和 DNA 甲基化模式间的精确关系也有待充分的说明（Tariq and Paszkowski 2004)。

组蛋白乙酰化是能被 HDAC 和 HAT 这两种相互作用进行动态调控的。与其相反，组蛋白甲基化被认为是目前为止最为牢固的表观遗传标志。然而，最近对于哺乳动物的研究识别了一种赖氨酸去甲基化酶（lysine demethylase，LSD1），它能去除 H3K4me1 和 H3K4me2 的甲基，却对 H3K4me3 没有作用（见第 10 章）。四个公认的 LSD 同源物在拟南芥中被编码，说明至少植物中某些组蛋白甲基化是可逆转的。

## 2.3　其他染色质蛋白

### 其他 Polycomb 蛋白

PcG 蛋白最初是在果蝇中作为维持同源异型基因抑制所需的因子而被识别的（见第 11 章）。动物中结构上分散的 PcG 蛋白以多蛋白复合体的形式协同作用抑制基因的表达。PRC1 复合物在植物和线虫中缺少，而在果蝇和哺乳动物中存在。PRC2 复合物被发现存在于植物和动物中，它通过 SET 结构域蛋白和 PcG E（Z）蛋白的组蛋白甲基化转移酶的活性主要完成 H3K27 的甲基化。

PRC2 核心组分的拟南芥同源物在为分析各种发育途径所设计的突变筛选实验中被发现。果蝇的 PRC2 组成由单拷贝基因编码。相反，拟南芥中编码这些蛋白的基因显示了至少 3 种 PRC2 复合物的功能多样性。三种复合物分别是：FIS(fertilization independent seeds)、EMF(embryonic flower) 和 VRN(vernalization)。这些复合物的靶基因特异性是不同的（Schubert et al. 2005；也见图 11.2）。

FIS 基因在无需受精而有部分种子发育的突变筛选实验中被发现，其主要靶蛋白是 MADS 盒转录因子 *PHERES*(Köhler et al. 2005)。EMF 复合物通常被认为抑制花的同源异型基因，例如 *Agamous* 和 *Apetala*3 基因。VRN 复合物成员之一 VRN2 被发现是基于它在春化实验中对表观遗传记忆的作用。种子促熟法打破了低温处理种子冬眠的过程。植物必须在适当的季节繁殖开花结果，通常他们都在低温时节延长阶段过后才开花。植物对冬季的记忆是表观遗传调控的，依赖 VRN2 蛋白，在较高温度下，VRN2 可以维持开花抑制因子 FLC 的编码基因处于转录抑制的状态，从而保证生长。*Vrn2* 的突变体中的 *FLC* 是缺失 H3K27me2 的，这与 PRC2 复合体在促进组蛋白甲基化方面的作用是一致的（Schubert et al. 2005）。

### 印记的组成

开花植物与哺乳动物是仅有的具有亲本印记的生物体（表 9-1）。亲本印记是一种表观遗传现象，即基因的表达量决定于它遗传的亲本。从双亲对立理论来看印记的进化（详见第 19 章），在开花植物和哺乳动物中亲本印记现象的存在可能反应了这样一个事实：这两类物种有特殊的母体组织为发育中的胚胎提供营养。在哺乳动物中，这个组织就是胎盘；而在植物中便是三倍体的胚乳，这种胚乳是在一定时期内出现的分化组织，包含一个父本和两个母本基因组（图 9-1）。其实，生物体的单基因的亲本印记现象最早是在玉米的胚乳中发现的（综述见 Alleman and Doctor 2000）。在拟南芥中，胚乳中表达的两个基因，*MEA* 及 *FWA*（*FWA* 是一种开花时间控制基因）就是有印记现象的。

在这些例子中，母本基因是激活的，可假定雌性配子体从 DME 催化的 CpG 主动去甲基化而致（见本章 2.1 节，活性 CpG 去甲基化和 DNA 糖苷水解），同时父本的基因则保持沉默（综述见 Autran et al. 2005）。令人好奇的是，虽然印记现象在植物和哺乳动物中是独立进化的，但是 DNA 甲基化和 PcG 蛋白却是两个过程中共有的元素（Köhler et al. 2005）。

染色质重塑蛋白

Switch2/Sucrose Non-Fermentable2(SWI2/SNF2) 染色质重塑因子组成了一类保守的 ATP 依赖型的染色质重塑子家族，这些重塑子可以排列核小体或者使组蛋白与 DNA 间的接触松开。遗传筛选只从拟南芥的基因组中筛到了大约 40 个 SWI2/SNF2 同源蛋白，这只是很少数的一部分（植物染色质数据库）。到目前为止，只有两种同源蛋白被证明调节 DNA 的甲基化——减少的 DNA 甲基化（decreased DNA methylation 1，DDM1）（Jeddeloh et al. 1999）及 RNA 指导的 DNA 甲基化的缺陷（defective in RNA-directed DNA methylation，DRD1）（Kanno et al. 2004）。*DDM1* 缺陷的突变体出现整个基因组 DNA 甲基化程度的降低，同时转录水平上重新激活了一些沉默的转座子和重复片段，这类突变体还呈现出发育和形态上的缺陷。不过这种缺陷只出现在纯合的 *ddm1* 植物数代培育之后，似乎是因为表观遗传突变的积累和由片段插入法插入的转座子基因的重新激活引起。DDM1 在哺乳动物中有一个同源蛋白，淋巴特异解旋酶（lymphoid-specific helicase，LSH），这种蛋白对整体意义上的 CpG 的甲基化及胚胎发育很重要。相反的，DRD1 是植物特有的，在 RdDM 中有特定的作用（见本章 3.4 节）。

除了解除特定重复片段的沉默外，几乎没有表观遗传上的改变是由修饰因子 MOM1(Morpheus' molecule) 导致的（Amedeo et al. 2000），*MOM1* 是一种植物特有的基因，含有不完整的 ATP 依赖的解旋酶结构。MOM 的作用与 DDM1/DNA 甲基化通路协同，却不依赖于该通路，表明植物中具有多层次的转录水平调控方式（Tariq and Paszkowski 2004）。

另外三种被公认的染色质重塑因子包括 SPD(Splayed)、PIE(Photoperiod-independent early flowering) 及 PKL(Pickle)，它们都是在其对应的缺陷突变体发育过程中发现的，但是还没有发现它们在特定染色质修饰中的作用（Wagner 2003）。

染色质组装因子

尽管 SWI2/SNF2 蛋白可能参与染色质的组装，染色质在复制后进行重组以及修复相关的 DNA 合成还需要其他的成分。染色体组装因子（chromatin assembly factor，CAF）由 3 个亚基组成，协助半组装的核小体进入复制叉。拟南芥中两个较大的 CAF 亚基（*fas1*，*fas2*）的突变体会出现异常形态（带化现象，fasciation，图 2f），产生 DNA 修复的缺陷以及重复片段的去抑制（Takeda et al. 2004）。这表明正确的核小体装卸是发育和表观遗传调控所必需的。同时发现 CAF 亚基的缺失不会影响 DNA 甲基化的维持，但是可以导致其他表观遗传标记的擦除，例如组蛋白修饰。CAF 第三个亚

基 MSI1 的低表达不会导致再次带化现象的出现，但是会导致畸形的种子发育及严重的形态上的改变（综述见 Hennig et al. 2005）。*BRU* 基因被认为不与任何已知的染色质组装蛋白相关，但是其的突变却和 *fas* 的突变体表型非常相似，于是被科学家认为非常有可能参与复制后期染色体组装过程中维持表观遗传信息和保证遗传信息的完整（Takeda et al. 2004）。最终，由于一种复制相关蛋白复合物的亚基 RPA2 的缺少，导致 DNA 易损伤且转录沉默减少，进而改变组蛋白修饰标记但并不改变 DNA 甲基化模式（Elmayan et al. 2005；Kapoor et al. 2005b）。

类异染色质蛋白

果蝇及哺乳动物中的 HP1(heterochromatin protein 1）以及它们在真菌中的同源蛋白是表达沉默的异染色质重要的组成部分。HP1 通过 chromo 结构域结合到甲基化的组蛋白 H3K9 上，协助传递基因沉默的状态，并形成异染色质型的结构域。拟南芥基因组编码一个与果蝇 HP1 同源的蛋白，被称为类异染色质蛋白（like heterochromatin protein，LHP1）（Gaudin et al. 2001）或 TFL2(terminal flower 2）（Kotake et al. 2003），它的突变体会改变植物的株形并改变叶的发育，引起早花。尽管这个突变体的表型表明它在植物基因表达中有重要的作用。但是似乎 LHP1 并不是像 HP1 那样，通过形成抑制型的染色质复合体作用的。拟南芥中的 LHP1 调控常染色质的位点，而常染色质并不是 DNA 甲基化的目标（Kotake et al. 2003；Tariq and Paszkowski 2004）。因此，植物中的 LHP1 发其他生物中的 HP1 应该是通过不同的方式进化而来的。

## 3. RNAi 介导的基因沉默通路的分子组成

传统的现代表观遗传研究关注 DNA 甲基化和组蛋白的修饰。在过去的几年中，有证据表明研究重心将转向 RNA 干扰（RNAi）通路指向的特异基因组区域。甚至，现在在很多真核生物不可能脱离 RNAi 机制分子组成来考虑表观遗传调控的问题（Matzke and Birchler 2005）。在植物中此说法尤其正确，因为植物中 RNAi 介导的基因沉默途径的存在远远超过了任何一种生物。

### 3.1　细述植物中 RNAi 介导的基因沉默

RNAi 和其相关的基因沉默体现了细胞对双链 RNA(dsRNA）的反应。在这些通路中，双链 RNA 被 RNaseIII 类似的内切核酸酶——Dicer 加工成小 RNA，这些小 RNA 决定了沉默的特异性，并特异性地结合到对应的目标核酸链上。小 RNA 整合形成多蛋白的沉默因子复合体，指导序列特异性的 mRNA 降解、抑制翻译（PTGS）或者指导染色质的修饰（TGS）。沉默因子复合体的一个重要成分就是 Argonaute 蛋白，它通过 PAZ 结构域结合到小 RNA 上。Argonaute 蛋白家族的成员组成了 RNAi 介导的基因沉默作用中最重要的一类蛋白家族，对于不同的沉默因子复合体具有不同的特异功能（综述见 Carmell et al. 2002）。除了参与病毒防御和转座子的调控，RNAi 介导的基因沉默在植物和动物的发育中有十分重要的作用。

RNAi介导的基因沉默的目的反映了它在进化中与复制时出现双链RNA的病原体是共同进化的，如RNA病毒和类病毒。甚至，它是与一种转基因——另一种外来核酸一同进化的，因而这些RNA病原体对于发现和研究植物中不同形式的RNA干扰介导的基因沉默非常有价值。

RNAi介导的基因沉默途径在植物中的多样性体现为：

(1) 编码RNAi核心组件的基因家族非常庞大而且功能多样：拟南芥基因组编码4个类似DICER(DCL)的蛋白和10个Argonaute(AGO)蛋白。

(2) 小RNA长度不同且功能多样，包括起源自转基因和病毒的21nt的短干扰RNA(siRNA)，还有几种内源的小RNA，例如21～24nt的微小RNA(microRNA)；21nt的反式作用的siRNA以及24～26nt的异染色质siRNA。

(3) 不同小RNA引起的不同的基因沉默模式：PTGS包含mRNA的降解或者翻译的抑制，TGS与表观遗传修饰相关，如DNA胞嘧啶的甲基化和组蛋白的甲基化。

(4) PTGS在抗病毒防御中有很重要的作用，它是由多种植物病毒蛋白在通路的不同阶段抑制基因沉默引起的。

(5) 某些过程的存在，如非细胞自主沉默(non-cell-autonomous silencing)和传递性(见本章3.2节，非细胞自主沉默和传递性)，依赖于RNA依赖的RNA聚合酶，而且这些酶中有6个由拟南芥基因组编码。

这些方面将在讲述植物RNA干扰介导的基因沉默的3条主要通路的框架时会提到(图9-3a～c)。然而，需要记住的是这些通路在很多点上是相互交叉的。已研究出功能的组件列在表9-2中。

## 3.2 通路一：转基因相关的转录后沉默及病毒诱导的基因沉默(PTGS/VIGS)

由转基因或病毒引起的RNAi介导的基因沉默似乎主要是宿主对于外来侵入的核酸、也包括病毒、转座子及转基因的主要防御方式。

### dsRNA的来源及加工

转基因载体可以正义、反义或者反向DNA重复序列的方式插入植物的基因组。病毒基因组可以是单链或双链的DNA或RNA。因此，在这条通路中，双链RNA可以通过很多途径产生。原则上，反义的转录产物可以与目标mRNA配对形成双链RNA。通过反向重复序列产生的翻译产物可以形成茎环结构的RNA。RNA病毒编码自己的依赖于RNA的RNA聚合酶(vRdRP)并通过双链RNA中间体完成复制，直接通过双链RNA进入到通路中。相反，正义的转基因和DNA病毒，如双生病毒(Geminivirus)，则需要细胞的依赖于RNA的RNA聚合酶RDR6以及其他的因子(SDE3、SGS3和WEX；表9-2)来完成双链RNA的合成。为了成为RDR6的底物，正义转基因的转录产物和DNA病毒被认为具有特殊的性质，如缺少5′帽子或者Poly(A)尾(综述见Meins et al. 2005)。

在PTGS通路中将双链RNA加工成siRNA所需要的DCL蛋白的活性至今仍未研

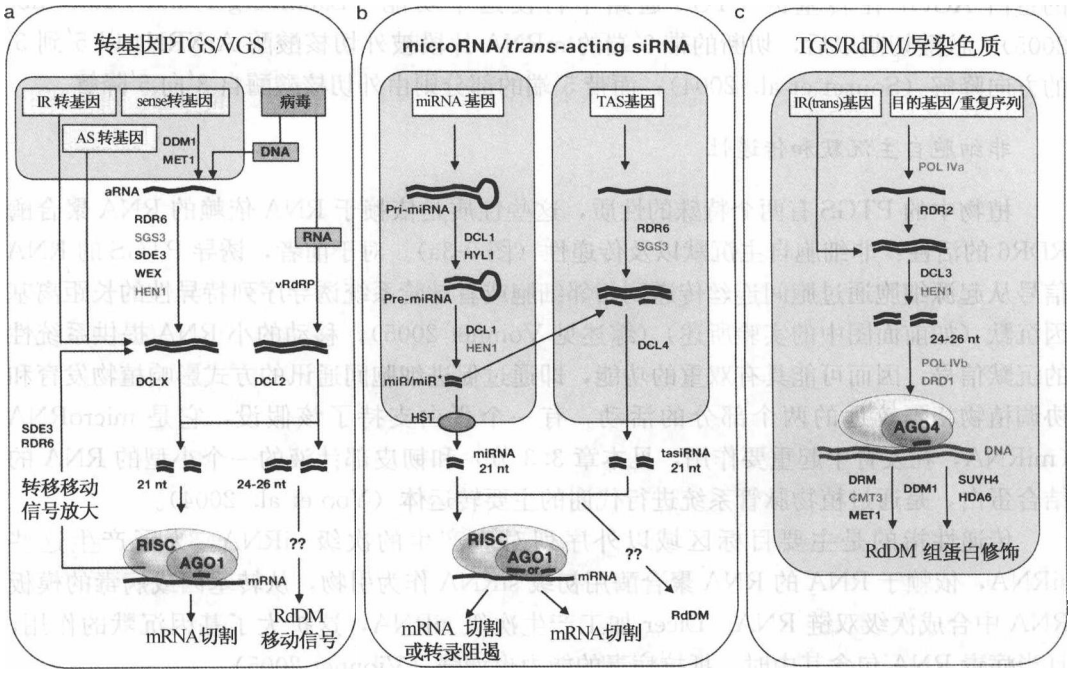



图 9-3　植物中 RNA 介导的基因沉默途径

尽管存在一些重叠和共有的组成，3 条主要途径在 dsRNA 来源、小 RNA 分类、靶序列属性以及引起的沉默模式等方面是不同的。含有 Argonaute 蛋白沉默效应复合物显示在亮灰色标记区内。黄色框内标记为已知在核内发生的过程。细节内容请参看正文，调节组分的名字请参考表 9-2。植物特异蛋白用绿色标记。PTGS，posttranscriptional gene silencing，转录后基因沉默；VIGS，virus-induced gene silencing，病毒诱导的基因沉默；TGS，transcriptional gene silencing，转录水平基因沉默；RdDM，RNA-directed DNA methylation，RNA 指导的 DNA 甲基化；IR，inverted repeats，反向重复；AS，antisense，反义；vRdRP，virally encoded RNA-dependent RNA polymerase，病毒编码的 RNA 依赖的 RNA 聚合酶；aRNA，aberrant RNA，异常 RNA；siRNA，short interfering RNA，短小干扰 RNA；RISC，RNA-induced silencing complex，RNA 诱导的沉默复合物（修改自 Meins et al. 2005）。

究清楚（DCLX）。通过 *dcl1* 部分丧失功能的突变体的测试表明，DCL1 似乎并不参与该加工过程。植物特异的蛋白 HEN1 在小 RNA 的 3′端的核苷酸碱基上加上甲基，起到保护它们不会尿嘧啶化并降解（Li et al. 2005）。DCL2 被证实参与通过一些但不是全部的 RNA 病毒形成 siRNA(Xie et al. 2004)。

PTGS 和 VIGS 导致产生两类不同大小的 siRNA（21～22nt 和 24～26nt），这两类 siRNA 有不同的功能（Baulcombe 2004）。通常，21nt 的 siRNA 指导 RNA 的剪切，而 24～26nt 的 siRNA 被称为异染色质 siRNA，指导同源 DNA 序列的表观遗传修饰（即 TGS；见本章 3.4 节）。

在 DCL 加工以后，siRNA 不再卷曲，其反义链与 Argonaute 蛋白家族的成员作用，组装成为 RNA 诱导的基因沉默复合体的一部分（RNA-induced silencing complex，RISC）。siRNA 指导的 RISC 在 siRNA-mRNA 形成的配对体的单位点对目标 mRNA 进行核苷酸的剪切。在哺乳动物中，剪切是由 Ago2 剪切子催化的（见第 8 章）。拟南芥

的蛋白 AGO1 在转基因 PTGS 通路中行使这个功能（Baumberger and Baulcombe 2005）。核酸切割以后，切断的带 3′端的 mRNA 片段被外切核酸酶 AtXRN4 以 5′到 3′的方向降解（Souret et al. 2004）；而带 5′端的部分则由外切核酸酶自 3′向 5′降解。

#### 非细胞自主沉默和传递性

植物中的 PTGS 有两个特殊的性质，这些性质是依赖于 RNA 依赖的 RNA 聚合酶 RDR6 的活性：非细胞自主沉默以及传递性（图 9-3a）。对于前者，诱导 PTGS 的 RNA 信号从起源细胞通过胞间连丝传递到相邻细胞或者脉管系统诱导序列特异性的长距离基因沉默（如前面图中的实验所述）（综述见 Voinnet 2005）。移动的小 RNA 提供系统性的沉默信号，因而可能具有双重的功能，即通过促进细胞间通讯的方式影响植物发育和协调植物相隔较远的两个部分的活动。有一个蛋白支持了该假设，它是 microRNA（miRNA，在发育中起重要作用，见本章 3.3 节）和韧皮部汁液的一个小型的 RNA 的结合蛋白，是通过植物脉管系统进行代谢的主要转运体（Yoo et al. 2004）。

传递性指的是主要目标区域以外序列对应产生的次级 siRNA。为了产生这些 siRNA，依赖于 RNA 的 RNA 聚合酶用初级 siRNA 作为引物，从转基因或病毒的模板 RNA 中合成次级双链 RNA。Dicer 加工产生次级 siRNA，这扩大了基因沉默的作用，且当病毒 RNA 包含其中时，抵抗病毒的能力也增强（Vionnet 2005）。

除了植物以外，同时发现非细胞自身基因沉默现象和转递性现象的只有线虫（*C. elegans*）（见第 8 章），因为它被公认有依赖于 RNA 的 RNA 聚合酶活性，而此活性在哺乳动物与果蝇中是不存在的。

#### 基因沉默的病毒抑制因子

植物病毒不仅诱导基因沉默和成为基因沉默的目标，它们同时编码可以抑制沉默作用的蛋白（综述见 Voinnet 2005）。这更支持了一个理论：PTGS 是对病毒的自然防御，因为这些蛋白形成了抗防御的病原体机制。大多植物病毒编码至少一种基因沉默抑制蛋白，该蛋白会在 PTGS 通路特定的步骤中发挥作用，尤其是双链 RNA 加工的下游。PTGS 通路被病毒抑制最好的证明是在杂色的大豆上，黑色斑点是一个色素基因受自然 PTGS 通路反式作用的结果（即再激活）（图 9-2e）（Senda et al. 2004）。最近还在一种昆虫病毒和哺乳动物逆转录病毒中发现了 RNA 干扰的病毒抑制现象（Lecellier et al. 2005；Vionnet 2005）。

### 3.3 通路二：通过 miRNA 和反式作用 siRNA 对植物发育的调控

在植物和动物中发现内源的 miRNA 为发育生物学和 RNA 干扰介导的基因沉默的研究开创了新的世纪（综述见 Bartel 2004）。miRNA 通过碱基配对到目标 mRNA 并诱导 mRNA 的切割或是抑制翻译来达到沉默基因表达的作用。miRNA 对植物发育的重要性体现为 miRNA 产生和基因沉默必须的很多基因——包括 *DCL1*、*AGO1*、*HEN1*、*HYL1* 及 *HST* 都在对发育突变体的筛选中被发现，后来又被发现对于 miRNA 的积累有很重要的作用。在这些蛋白缺陷的突变体的表型上可以揭示 miRNA 的很多功能：分

生组织功能、器官极性、脉管发育、花型以及应激作用/激素应答（综述见 Kidner and Martenssen 2005）。最近发现 miRNA 参与生成一种新的小 RNA——反式作用的 siRNA。

### miRNA 的功能和生物学形成

miRNA 初始发现于克隆小片段的小 RNA，这些小 RNA 的长度从 18～28nt 不等。在植物中，miRNA 与目标 mRNA 高度的互补性给通过计算方式识别其余的 miRNA 带来很大的便利。到目前为止，拟南芥中编码 27 个不同 mRNA 的 92 个位点被发现，在小鼠中发现的数目是相似的。其中很多 miRNA 基因的表达是随着发育或者环境调控的。拟南芥中 50%的 miRNA 的目标是转录因子，它们其中很多被认为是分生组织形成和功能的调控因子，这些功能是在确定它们是 miRNA 目标之前就已经确定的。相反的是，动物细胞中 miRNA 并不以转录因子为目标，而是调控细胞很多层面上的不同基因。拟南芥 miRNA 通路有两个非常重要的蛋白：DCL1 和 AGO1，它们本身也被 miRNA 调控的，是一种负反馈调节的途径（Kidner and Martienssen 2005）。

很多 miRNA 在真核动物的进化中是保守的（Axtell and Bartel 2005），有些甚至很古老。一个代表性的例子，在开花植物、裸子植物和更原始的植物中，调控分生组织形成和侧向器官不对称等转录因子的编码 mRNA 就与同类的 miRNA 完美匹配。这表明该功能的保守性至少维持了 4 亿年（Floid and Bowman 2004）。

miRNA 的编码区通常位于两个编码蛋白的基因之间或者位于内含子内。它们源于不完整的 RNA 茎环结构前体，这些前体长度为 70～300 核苷酸不等，并且是由依赖 DNA 的 RNA 聚合酶Ⅱ转录完成的。植物的 miRNA 前体的加工过程在核中完成，并需要经过多个步骤。首先，核中 DCL1 去除前体 miRNA 的两端。该步骤需要双链 RNA 结合蛋白 HYL1 的协助，HYL1 蛋白是最初在其突变体的激素应答缺陷表型中发现的（Han et al. 2004；Vasquez et al. 2004a）。第二步包括通过 DCL1 释放双链 miRNA（miR/miR*，图 9-3b），HEN1 对其 3′端进行甲基化。将 miR/miR* 从核运送到胞质需要 HASTY 蛋白（HST），HASTY 是哺乳动物 Exportin 5 的同源蛋白（Park et al. 2005）。在核成分中也发现了成熟的 miRNA，表明有些 miRNA 可能留在核中指导表观遗传修饰。甚至，一些互补于剪切且新生转录因子的转录产物的 miRNA 会诱导目标基因下游的 DNA 序列发生胞嘧啶甲基化，但是其机理尚未得知（Schubert et al. 2005）。

植物中 miRNA 的来源与哺乳动物中有所不同，哺乳动物细胞质中含有单个 Dicer 酶，而核中含有第二个 RNA 酶Ⅲ的活性的 Drosha 蛋白。Drosha 与双链 RNA 结合蛋白 Pasha 一起剪除初级转录产物（pri-miRNA）的末端，而 Drosha 或者 Pasha 在植物中都不含有同源蛋白。产生的前体 miRNA（pre-miRNA）通过蛋白 Exportin 5 介导的通路转运到细胞质中，最后由 Dicer 完成对成熟 miRNA 的加工（Du and Zamore 2005；Kim 2005）。

### 植物 miRNA 指导 mRNA 的剪切

通常，动物的 miRNA 与其靶目标 mRNA 的互补并不完美，主要通过结合 3′-UTR

区域的多个位点实现对翻译的抑制。相反的是，与目标区域完美匹配的植物 miRNA 还对 mRNA 的剪切有协助作用，与 siRNA 的作用机制类似。但是随着研究的继续，这两条“规则”的例外越来越多。例如，植物 miR172 可以抑制翻译，而特定的哺乳动物的 miRNA 也可以导致 mRNA 被剪切（综述见 Du and Zamore 2005）。

AGO1 是 Argonaute 蛋白家族最重要的成员，也是拟南芥中 miRNA 指导的 RISC 中行使 mRNA 切割功能的组件（Baumberger and Baulcombe 2005）。AGO1 的发现早于 miRNA 的发现，是从叶子发育有缺陷的突变体中筛选得到的。Argonaute 这个名字源于 *agol* 突变体的表型，这种突变体很像钓钩，因为其叶子细且窄。*Agol* 的突变体呈现顶端分生组织的缺陷，这与 PNH/ZLL/AGO10 突变体的表型相似（表 9-2），AGO10 与 AGO1 非常相似，但似乎并不是 PTGS 通路需要的（Vaucheret et al. 2004）。AGO 蛋白在植物分生组织中的重要功能与它们在干细胞维持方面的功能是一致的（Carmell et al. 2002；Kidner and Martienssen 2005）。

反式作用的 siRNA

反式作用（trans-acting）的内源 siRNA(ta-siRNA) 是最近在拟南芥中发现的一种新的小 RNA。ta-siRNA 可以诱导其目标 mRNA 的切割，它同时具有 siRNA 和 miRNA 的特点。与 siRNA 相似的是，ta-siRNA 的双链 RNA 的合成依赖于 RDR6 及 SGS3；而与 miRNA 相似的是，ta-siRNA 起源于基因组的区域，且与它的目标 mRNA 序列很少雷同。为了保证能与目标 mRNA 配对的正确 siRNA 的形成，miRNA 通过 DCL4 酶完成了 dsRNA 前体的逐步剪切 RNA 前体（图 9-3b）（Allen et al. 2005；Gascoilli et al. 2005）。

ta-siRNA 在发育时间调控上具有功能。在发育过程中，开花植物的发育过程经历 2 个相变：植物相变，包含了从幼体到成体的转变以及繁殖相变，产生带有花苞的枝条而不是植物性的发芽（图 9-1）。尽管对开花诱导的遗传分析已经取得了很大的进展，但是对于植物期的相变了解依旧很少。一种 Argonaute 蛋白，ZIPPY/AGO7，在这个转化过程中有特定的作用。对出现早熟植物相变现象的突变体的筛选发现了 PTGS 中两个重要的基因，*RDR6* 及 *SGS3*，而该突变体与 *zip/ago7* 突变体非常相似（图 9-3b）（Peregrine et al. 2004）。进一步的分析表明在 *rdr6* 和 *sgs3* 突变体中被上调表达的几个基因通过转录后调节被 ta-siRNA 所沉默（Vazquez et al. 2004b）。这些发现表明 PTGS 机制的组件不只是在病毒防御和转基因沉默上具有重要功能，也同时对发育转变的时间具有调控功能。现在还未知 ta-siRNA 在动物中是否有对应的形式存在。

## 3.4 通路三：转基因相关的转录水平沉默，RNA 指导的 DNA 甲基化和异染色质的形成

现今 RNA 干扰介导的转录水平基因沉默的概念是源于早期对植物中同源性依赖的基因沉默的研究，该基因沉默是由多拷贝的启动子区引起的，正源于植物中 RNA 指导的 DNA 甲基化的研究（综述见 Matzke and Matzke 2004）。越来越多的研究开始关注裂殖酵母中对 RNA 干扰介导的异染色质形成以及哺乳动物细胞中 siRNA 介导的 TGS

(见第 6、8 章)，这些研究拓展了这种行为在动植物中的范围，也再次确认了 RNA 干扰机制的普遍存在。

RNA 指导的 DNA 甲基化

RdDM 最初是在被类病毒感染的烟草植物中发现的（Wassenegger et al. 1994)。类病毒是十分微小的植物病原体，仅仅包含一个不编码蛋白的，几百碱基的环状 RNA。在以前的实验中，发现复制中的类病毒可以引起类病毒 cDNA 的从头甲基化，而类病毒的 cDNA 会整合到烟草的基因组中。随后，利用转基因系统发现 RdDM 需要双链 RNA，这些双链 RNA 可以被加工成小 RNA，这正是 RNA 干扰的标志。RNA 病毒是唯一被发现在细胞质中复制的病毒，RNA 病毒也可以引起同源性的核 DNA 的甲基化，表明 PTGS 产生的细胞质中的小 RNA 可以进入细胞核中并诱导表观遗传变化。包含启动子序列的双链 RNA 可以指导 DNA 甲基化以及同类目标启动子的转录水平沉默（综述见 Mathieu and Bender 2004；Matzke et al. 2005)。

在植物中，RNA 会诱导一种特殊的从头甲基化模式，这种模式的特点是在全序列范围内对胞嘧啶进行修饰，大部分位于 RNA-DNA 序列一致区。这个特点暗示 RNA-DNA 碱基的互补为从头甲基化提供了反应的底物，但是仍旧需要实验来确认。同时，不对称的 CpNpN 甲基化的维持在缺少了引发物 RNA 后效率很低，而对称的 CpG 和 CpNpG 甲基化即使 RNA 在不同启动子上，也可以不同程度地维持其甲基化的水平。甲基化维持效率上的不同可能反映了序列组成上的不同或者是组蛋白修饰模式的差异。

结合转基因和内源基因系统基因筛选得到的综合数据可提示 RdDM 和 TGS 所需要的分子成分。转基因系统依赖于已转录的反向重复序列或病毒能产生与目标 DNA 位点具有同源性的 dsRNA。内源基因在正向遗传筛选中提供了丰富的信息，例如氨基苯甲酸磷酸核糖转移异构酶（phosphoribosyl anthranilate isomerase，PAI）基因家族以及 *SUPERMAN* 基因（Chan et al. 2005)。这些基因具有一定的特点，使得他们可以成为 RdDM 或 TGS 的目标或诱导物。例如，*PAI* 基因家族包含 4 个成员，其中 2 个具有反向重复排列的特点。从不相关的上游启动子转录反向重复序列会得到一段双链 RNA，这个双链 RNA 以 *PAI* 基因单拷贝为目标进行甲基化和沉默。

植物特异的 RdDM 机制

多数情况下，保守的 DNA 甲基转移酶和组蛋白修饰酶是 RdDM 所必需的（本章 2.1 和 2.2 节)。所有序列胞嘧啶的从头甲基化是由保守 DRM 型的 DNA 甲基转移酶所催化的。保守的 MET1 和植物特有的 CMT3 主要的功能是维持 CpG 和 CpNpG 核酸组的甲基化，尽管其被报道对从头甲基化的作用很小。保守的组蛋白去乙酰化酶 HDA6 和 SWI2/SNF2 蛋白 DDM1 协助某些位点 CpG 甲基化的维持。组蛋白甲基转移酶 KYP/SUVH4 参与特定位点由 RNA 诱导的 CpNpG 甲基化的维持（综述见 Chan et al. 2005)。

最近一项令人惊奇的发现表明 RdDM 需要一种植物特异的被称作“Pol Ⅳ”的 RNA 聚合酶的参与。目前为止在所有的真核生物中有 3 种依赖 DNA 的 RNA 聚合

酶——Pol Ⅰ，Pol Ⅱ，Pol Ⅲ，均包含有多个由不同基因编码的亚基。Pol Ⅳ存在的第一个依据是对拟南芥基因组序列的分析，其显示了有些基因编码最大和第二大的某种非典型 RNA 聚合酶的亚基，而这种聚合酶是植物特有的。看起来有两种功能上不同的 Pol Ⅳ复合体，特点是两种 Pol Ⅳ的最大亚基不同，但都与一个共同的亚基（第二大）相互作用。Pol Ⅳa 是产生 siRNA 必需的，这是由起始转录靶基因的功能推测而来的（Herr et al. 2005；Onodera et al. 2005）。最初的转录产物被 RDR2 作为模板产生双链 RNA，然后由 DCL3 加工，DCL3 具有核内活性可以通过转座子和重复序列产生 24 个核苷酸的异染色质 siRNA(Xie et al. 2004)。在 siRNA 产生的下游，Pol Ⅳb 与植物特异的类 SWI2/SNF2 蛋白 DRD1 相互作用，提供 DNA 甲基化信号（Kanno et al. 2005），并可能与 AGO4 协同效应（图 9-3c）(Chen et al. 2005)。然而 Pol Ⅳb 是否参与转录 RNA 还未知，但是它的净效应是形成异染色质结构，以便允许 DNA 甲基转移酶在 siRNA 目标位点催化胞嘧啶的从头甲基化。尽管其他真核生物不含有 Pol Ⅳ亚基，但是负责转录 mRNA 前体的 Pol Ⅱ的两个亚基是裂殖酵母中 RNA 干扰介导的异染色质形成所必需的（见第 6、8 章）。

尽管人类细胞中启动子指导的 siRNA 能诱导 TGS，仍有相反的报道质疑是否 TGS 伴随可检测的 DNA 甲基化的存在（Kawasaki et al. 2005；Ting et al. 2005）。植物中很多 RdDM 所需的蛋白仅是在植物界中发现（图 9-3c）。因此，如果 RdDM 在哺乳动物中发生，则其机制或蛋白系统与植物中的应有差别。

### 通过 RNA 干扰介导的 TGS 内源基因的沉默

很多转座子和 DNA 重复序列如 5S rRNA 和拟南芥 4 号染色体上富含转座子的异染色质纽都通过 RNA 介导的机制实现转录沉默和甲基化（Lippman and Martienssen 2004；Chan et al. 2005）。内源的 DNA 目标反应了 RNA 介导的 TGS 通路的自然功能，这种功能体现为抑制转座子活动以及将重复序列引入异染色质部分。然而，包含转座子插入序列的植物基因自身可以成为 RNA 干扰介导的沉默和甲基化的目标序列。例如，拟南芥成花基因 *FWA* 启动子中有从转座子中衍生而来的重复序列，这些序列就是同源 siRNA 甲基化的目标（Lippman and Martienssen 2004），因此可以沉默那些在营养组织中不被需要的基因。在某些拟南芥品种中，具有抑制成花作用的 *FLC* 基因内含子中的 *Mu* 元件使基因对产生于分散的 *Mu* 拷贝的 siRNA 介导的抑制性染色质修饰的易感性增强（Liu et al. 2004）。*FLC* 的低表达可以加速开花时间，在特定的环境下具有适应的显著性。由于很多植物基因在启动子或内含子的间歇区具有转座子插入序列，这种模式的调控可能是植物界中常见的。甚至，随着对表观遗传调控了解逐步加深，McClintock 关于转座子作为调控元件调控宿主基因和发育的观点得到了越来越多的支持（McClintock，1956）。

## 4. 没有 RNA 参与的表观遗传调控

尽管小 RNA 提供了特异性，但是却不能涵盖植物中所有表观遗传修饰的模式。例

如，MOM 是一种带有部分 SNF2 结构域的蛋白，但没有发现参与 RNA 干扰介导的 TGS 通路。同时也没有证据表明蛋白中的 PcG 蛋白是通过小 RNA 找到他们的目标序列的。其他形式的信号，例如非转录重复序列同源配对或特殊的序列组成，也可能参与异染色质的形成和募集 DNA 甲基转移酶。例如，RNA 干扰的机理在链孢霉的 DNA 甲基化和组蛋白甲基化中是可有可无的，而富含 TA 的片段则是修饰作用更偏向的位点（见第 6 章）。此外，在裂殖酵母中有些异染色质的形成途径是不依赖于 RNA 干扰的（见第 6、8 章）。

植物中有一种不寻常的、不涉及 RNA 干扰的表观遗传现象被称为副突变（paramutation）。当某些被称作副突变基因的等位基因施加表观遗传印记于易感等位基因（susceptible allele）（亦称副突变等位基因，paramutable allele）时，副突变就发生了。表观遗传印记是通过减数分裂遗传给后代，即便在两个相互作用的等位基因分离时仍旧持续存在。副突变是孟德尔遗传定律的一个例外，孟德尔定律认为等位基因从杂合体中分离后是保持不变的。副突变的现象在数十年前就在玉米和土豆中发现，但其中的机理始终为谜。副突变类似的现象最近在哺乳动物中发现，表明该现象不是仅局限于植物界（综述见 Chandler and Stam 2004）。玉米中的 B 基因座是研究最多的副突变的例子，B 基因座包含一系列的正向重复序列，大约位于转录起始点 100kb 的位置，具有介导副突变的功能，但是机理未知。尽管没有完全排除 RNA 介导的基因沉默的可能，但是依赖于等位基因配对的其他的机理仍列于考虑中（综述见 Stam and Mittelsten Scheid 2005）。

## 5. 展　望

在本章中，我们讨论了植物有关的基本表观遗传原理以及它们与其他生物表观遗传调控现象的关系。很显然，植物与其他生物共同拥有许多表观遗传调控的特征，但是它们也进化产生大量植物特有的变化和创新。这显示出了植物发育独特的一面，也解释了他们在不可预知的环境中超强的存活和繁殖能力。

植物特有的表观遗传调控中最显著的一点就是用于回复性表观遗传修饰的嵌入式系统，该系统对植物发育的可塑性和适应性贡献巨大。植物具有在未分化细胞中诱导或者消除抑制性修饰的能力，诱导通过 RdDM 和组蛋白修饰完成，而消除则通过 DNA 糖苷酶完成（如 DME 和 ROS1），这种能力使植物可以在不涉及 DNA 复制周期的情况下启动表观遗传的修饰作用。表观遗传标记从基因组的易消除性或许可以解释植物从一个体细胞克隆出整株植物的相对简便性。尽管如此，表观遗传标记的诱导或者消除在个体植物中是无法做到完美或者一致的，因而就形成了遗传意义上说应该完全一致的克隆植株却出现表观遗传的差异性。这种体细胞克隆的差异性可以在植物繁殖过程中体现出来。

相似的是，表观遗传标记在有性生殖过程的分化遗传可以在自然植物群落中产生表观遗传差异性。可以通过固定某个具有适应显著性的表观遗传等位基因来筛选不同的植物。就如我们介绍过的春化作用一样，环境因素可以引发植物中的表观遗传修饰，改变

生理应答。因此，当环境或者刺激引起的表观遗传修饰本身具有适应性且能够从分生组织进入干细胞的话，植物对这些环境、刺激便具有记忆能力。在特定的情况下，植物自发进行表观遗传的修饰，而在另一些情况下，植物实行程序化的表观遗传修饰，因而对全部这些条件的定义可以对这些修饰的生物功能的阐述有更大的帮助。同样的，如果植物表观遗传标记的减数分裂遗传的机理能得到解析的话，科学家就可以利用此机理提高花卉培养和农业培养的质量。

除了对环境刺激的合适应答，植物在进化和繁殖历史中面临了多种多样的基因组挑战。多倍性或二倍性对于由于不明过程的基因组刺激（genome shock）引起的表观遗传修饰上有重要的作用，基因组刺激（genome shock）是对于非正常刺激的一种应答，这种非正常的刺激会导致转座子大范围的移动（McClintock 1983）。杂种优势的原因，即杂合体相比较野生的亲本系而言有优势的存活能力，仍然未知，但是似乎与组合两个相关但不同的基因组引起的表观遗传变化有关。相似的是，多倍体将整个基因组合并或者加倍，对于表观遗传变化而言有数不尽的可能性。对植物多倍体表观遗传结果的认识也可以帮助了解我们自己的进化历史，因为人类进化史也被认为是经历了两次整个基因组的复制（Forlong and Holland 2004）。显然，在这个层面上，植物的表观遗传学可以为人类生物学的研究提供丰富的信息，我们甚至可以称它们为“表观遗传调控的大师”。

赵 璟 王大亮 译
王大亮 校
孙方霖 审校

## 参考文献


Adams K.L. and Wendel J.F. 2005. Polyploidy and genome evolution in plants. *Curr. Opin. Plant Biol.* **8:** 135–141.

Alleman M. and Doctor J. 2000. Genomic imprinting in plants: Observations and evolutionary implications. *Plant Mol. Biol.* **43:** 147–161.

Allen E., Xie Z., Gustafson A.M., and Carrington J.C. 2005. microRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. *Cell* **121:** 207–221.

Amedeo P., Habu Y., Afsar K., Mittelsten Scheid O., and Paszkowski J. 2000. Disruption of the plant gene MOM releases transcriptional silencing of methylated genes. *Nature* **405:** 203–206.

Autran D., Huanca-Mamani W., and Vielle-Calzada J.P. 2005. Genomic imprinting in plants: The epigenetic version of an Oedipus complex. *Curr. Opin. Plant Biol.* **8:** 19–25.

Axtell M.J. and Bartel D.P. 2005. Antiquity of microRNAs and their targets in land plants. *Plant Cell* **17:** 1658–1673.

Bartel D.P. 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116:** 281–297.

Baulcombe D.C. 2004. RNA silencing in plants. *Nature* **431:** 356–363.

Baumberger N. and Baulcombe D.C. 2005. *Arabidopsis* ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. *Proc. Natl. Acad. Sci.* **102:** 11928–11933.

Baumbusch L.O., Thorstensen T., Krauss V., Fischer A., Naumann K., Assalkhou R., Schulz I., Reuter G., and Aalen R.B. 2001. The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res.* **29:** 4319–4333.

Burch-Smith T.M., Anderson J.C., Martin G.B., and Dinesh-Kumar S.P. 2004. Applications and advantages of virus-induced gene silencing for gene function studies in plants. *Plant J.* **39:** 734–746.

Carmell M.A., Xuan Z., Zhang M.Q., and Hannon G.J. 2002. The Argonaute family: Tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev.* **16:** 2733–2742.

Chan S.W.-L., Henderson I.R., and Jacobsen S.E. 2005. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* **6:** 351–360.

Chandler V.L., and Stam M. 2004. Chromatin conversations: Mechanisms and implications of paramutation. *Nat. Rev. Genet.* **5:** 532–544.

Chandler V.L., Eggleston W.B., and Dorweiler J.E. 2000. Paramutation in maize. *Plant Mol. Biol.* **43:** 121–145.

Choi Y., Gehring M., Johnson L., Hannon M., Harada J.J., Goldberg R.B., Jacobsen S.E., and Fischer R.L. 2002. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* **110:** 33–42.

Cubas P., Vincent C., and Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401:** 157–161.

Du T. and Zamore P.D. 2005. microRNAPrimer: The biogenesis and function of microRNA. *Development* **132:** 4645–4652.

Elmayan T., Proux F., and Vaucheret H. 2005. *Arabidopsis* RPA2: A genetic link among transcriptional gene silencing, DNA repair, and DNA replication. *Curr. Biol.* **15:** 1919–1925.

Fedoroff N.V. and Chandler V. 1994. Inactivation of maize transposable elements. In *Homologous recombination and gene silencing in plants* (ed J. Paszkowski), pp. 349-385. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Floyd S.K. and Bowman J.L. 2004. Gene regulation: Ancient microRNA target sequences in plants. *Nature* **428:** 485–486.

Furlong R.F. and Holland P.W. 2004. Polyploidy in vertebrate ancestry: Ohno and beyond. *Biol. J. Linnean Soc.* **82:** 425–430.

Gasciolli V., Mallory A.C., Bartel D.P., and Vaucheret H. 2005. Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing *trans*-acting siRNAs. *Curr. Biol.* **15:** 1494–1500.

Gaudin V., Libault M., Pouteau S., Juul T., Zhao G., Lefebvre D., and Grandjean O. 2001. Mutations in *LIKE HETEROCHROMATIN PROTEIN 1* affect flowering time and plant architecture in *Arabidopsis*. *Development* **128:** 4847–4858.

Gong Z., Morales-Ruiz T., Ariza R.R., Roldan-Arjona T., David L., and Zhu J.K. 2002. *ROS1*, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell* **111:** 803–814.

Han M.H., Goud S., Song L., and Fedoroff N. 2004. The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc. Natl. Acad. Sci.* **101:** 1093–1098.

Heitz E. 1928. Das Heterochromatin der Moose. *I. Jahrb. Wiss. Bot.* **69:** 762–818.

Hennig L., Bouveret R., and Gruissem W. 2005. MSI1-like proteins: An escort service for chromatin assembly and remodeling complexes. *Trends Cell Biol.* **15:** 295–302.

Herr A.J., Jensen M.B., Dalmay T., and Baulcombe D.C. 2005. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308:** 118–120.

Hsieh T.-F. and Fischer R.L. 2005. Biology of chromatin dynamics. *Annu. Rev. Plant Biol.* **56:** 327–351.

Hung M.-S. and Shen C.-K. 2003. Eukaryotic methyl-CpG-binding domain proteins and chromatin modification. *Eukaryot. Cell* **2:** 841–846.

Jackson J.P., Lindroth A.M., Cao X., and Jacobsen S.E. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416:** 556–556.

Jeddeloh J.A., Stokes T.L., and Richards E.J. 1999. Maintenance of genomic methylation requires a SWI2/SNFs-like protein. *Nat. Genet.* **22:** 94–97.

Jorgensen R.A. 2003. Sense cosuppression in plants: Past, present, and future. In *RNAi: A guide to gene silencing* (ed. G.J. Hannon), pp. 5–22. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Kaeppler S.M., Kaeppler H.F., and Rhee Y. 2000. Epigenetic aspects of somaclonal variation in plants. *Plant Mol. Biol.* **43:** 59–68.

Kanno T., Mette M.F., Kreil D.P., Aufsatz W., Matzke M., and Matzke A.J. 2004. Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr. Biol.* **14:** 801–805.

Kanno T., Huettel B., Mette M.F., Aufsatz W., Jaligot E., Daxinger L., Kreil D.P., Matzke M., and Matzke A.J.M. 2005. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat. Genet.* **37:** 761–765.

Kapoor A., Agius F., and Zhu J.K. 2005a. Preventing transcriptional gene silencing by active DNA demethylation. *FEBS Lett.* **579:** 5889–5898.

Kapoor A., Agarwal M., Andreucci A., Zheng X., Gong Z., Hasegawa P.M., Bressan R.A., and Zhu J.K. 2005b. Mutations in a conserved replication protein suppress transcriptional gene silencing in a DNA-methylation-independent manner in *Arabidopsis*. *Curr. Biol.* **15:** 1912–1918.

Kawasaki H., Taira K., and Morris K.V. 2005. siRNA induced transcriptional gene silencing in mammalian cells. *Cell Cycle* **4:** 442–448.

Kidner C.A. and Martienssen R.A. 2005. The developmental role of microRNA in plants. *Curr. Opin. Plant Biol.* **8:** 1–7.

Kim V.N. 2005. MicroRNA biogenesis: Coordinated cropping and dicing. *Nat. Rev. Genet.* **6:** 376–385.

Köhler C., Page D.R., Gagliardini V., and Grossniklaus U. 2005. The *Arabidopsis thaliana* MEDEA polycomb group protein controls expression of PHERES1 by parental imprinting. *Nat. Genet.* **37:** 28–30.

Kotake T., Takada S., Nakahigashi K., Ohto M., and Goto K. 2003. *Arabidopsis TERMINAL FLOWER 2* gene encodes a heterochromatin protein 1 homolog and represses both *FLOWERING LOCUS T* to regulate flowering time and several floral homeotic genes. *Plant Cell Physiol.* **44:** 555–564.

Kress C., Thomassin H., and Grange T. 2001. Local DNA demethylation in vertebrates: How could it be performed and targeted? *FEBS Lett.* **494:** 135–140.

Lecellier C.-H., Dunoyer P., Arar K., Lehmann-Che J., Eyquem S., Himber C., Saib A., and Voinnet O. 2005. A cellular microRNA mediates antiviral defense in human cells. *Science* **308:** 557–560.

Li J., Yang Z., Yu B., Liu J., and Chen X. 2005. Methylation protects miRNAs and siRNAs from 3′-end uridylation activity in *Arabidopsis*. *Curr. Biol.* **15:** 1501–1507.

Lippman Z. and Martienssen R. 2004. The role of RNA interference in heterochromatic silencing. *Nature* **431:** 364–370.

Liu J., He Y., Amasino R., and Chen X. 2004. siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis*. *Genes Dev.* **18:** 2873–2878.

Loidl P. 2004. A plant dialect of the histone language. *Trends Plant Sci.* **9:** 84–90.

Malagnac F., Bartee L., and Bender J. 2002. An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J.* **21:** 6842–6852.

Mathieu O. and Bender J. 2004. RNA-directed DNA methylation. *J. Cell Sci.* **117:** 4881–4888.

Matzke M.A. and Birchler J.A. 2005. RNAi-mediated pathways in the nucleus. *Nat. Rev. Genet.* **6:** 24–35.

Matzke M. and Matzke A.J.M. 2004. Planting the seeds of a new paradigm. *PLoS Biol.* **2:** 528–586.

Matzke M., Kanno T., Huettel B., Jaligot E., Mette M.F., Kreil D.P., Daxinger L., Rovina P., Aufsatz W., and Matzke A.J.M. 2005. RNA-directed DNA methylation. In *Plant epigenetics* (ed. P. Meyer), pp. 69–96. Blackwell Publishing, Oxford, United Kingdom.

McClintock B. 1956. Intranuclear systems controlling gene action and mutation. *Brookhaven Symp. Biol.* **8:** 58–74.

———. 1983. The significance of responses of the genome to challenge. Nobel lecture. http://nobelprize.org/medicine/laureates/1983/mcclintock-lecture.pdf

Meins F., Si-Ammour A., and Blevins T. 2005. RNA silencing systems and their relevance to plant development. *Annu. Rev. Cell Dev. Biol.* **21:** 297–318.

Naumann K., Fischer A., Hofmann I., Krauss V., Phalke S., Irmler K., Hause G., Aurich A.-C., Dorn R., Jenuwein T., and Reuter G. 2005. Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J.* **24:** 1418–1429.

Onodera Y., Haag J.R., Ream T., Nunes P.C., Pontes O., and Pikaard C. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120:** 613–622.

Pandey R., Müller A., Napoli C.A., Selinger D.A., Pikaard C.S., Richards E.J., Bender J., Mount D.W., and Jorgensen R.A. 2002. Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. *Nucl. Acids Res.* **30:** 5036–5055.

Park M.Y., Wu G., Gonzalez-Sulser A., Vaucheret H., and Poethig R.S. 2005. Nuclear processing and export of microRNAs in *Arabidopsis*. *Proc. Natl. Acad. Sci.* **102:** 3691–3696.

Peregrine A., Yoshikawa M., Wu G., Albrecht H.L., and Poethig R.S. 2004. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of *trans*-acting siRNAs in *Arabidopsis*. *Genes. Dev.* **18:** 2368–2379.

Pikaard C.S. 2000. The epigenetics of nucleolar dominance. *Trends Genet.* **16:** 495–500.

Rocha P.S., Sheikh M., Melchiorre R., Fagard M., Boutet S., Loach R., Moffatt B., Wagner C., Vaucheret H., and Furner I. 2005. The *Arabidopsis HOMOLOGY-DEPENDENT GENE SILENCING1* gene codes for an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent silencing. *Plant Cell* **17:** 404–417.

Schubert D., Clarenz O., and Goodrich J. 2005. Epigenetic control of plant development by Polycomb-group proteins. *Curr. Opin. Plant Biol.* **8:** 553–561.

Senda M., Masuta C., Ohnishi S., Goto K., Kasai A., Sano T., Hong J.-S., and MacFarlane S. 2004. Patterning of virus-infected *Glycine max* seed coat is associated with suppression of endogenous silencing of chalcone synthase genes. *Plant Cell* **16:** 807–818.

Souret F.F., Kastenmayer J.P., and Green P.J. 2004. AtXRN4 degrades mRNA in *Arabidopsis* and its substrates include selected miRNA targets. *Mol. Cell* **15:** 173–183.

Springer N.M. and Kaeppler S.M. 2005. Evolutionary divergence of monocot and dicot methyl-CpG-binding domain proteins. *Plant Physiol.* **138:** 92–104.

Springer N.M., Napoli C.A., Selinger D.A., Pandey R., Cone K.C., Chandler V.L., Kaeppler H.F., and Kaeppler S.M. 2003. Comparative analysis of SET domain proteins in maize and *Arabidopsis* reveals multiple duplications preceding the divergence of monocots and dicots. *Plant Physiol.* **132:** 907–925.

Stam M. and Mittelsten Scheid O. 2005. Paramutation: An encounter leaving a lasting impression. *Trends Plant Sci.* **10:** 283–290.

Takeda S., Tadele Z., Hofmann I., Probst A.V., Angelis K.J., Kaya H., Araki T., Mengiste T., Mittelsten Scheid O., Shibahara K., et al. 2004. BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in *Arabidopsis*. *Genes Dev.* **18:** 782–793.

Tariq M. and Paszkowski J. 2004. DNA and histone methylation in plants. *Trends Genet.* **20:** 244–251.

Ting A.H., Schuebel K.E., Herman J.G., and Baylin S.B. 2005. Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. *Nat. Genet.* **37:** 906–910.

Vaucheret H., Vazquez F., Crété P., and Bartel D.P. 2004. The action of *ARGONAUTE1* in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev.* **18:** 1187–1197.

Vazquez F., Gasciolli V., Crete P., and Vaucheret H. 2004a. The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr. Biol.* **14:** 346–345.

Vazquez F., Vaucheret H., Rajagopalan R., Lepers C., Gasciolli V., Mallory A.C., Hilbert J.-L., Bartel D.P., and Crété P. 2004b. Endogenous *trans*-acting siRNAs regulate the accumulation of *Arabidopsis* mRNAs. *Mol. Cell* **16:** 69–79.

Voinnet O. 2005. Induction and suppression of RNA silencing: Insights from viral infections. *Nat. Rev. Genet.* **6:** 206–220.

Wagner D. 2003. Chromatin regulation of plant development. *Curr. Opin. Plant Biol.* **6:** 20–28.

Wang X.J., Gaasterland T., and Chua N.H. 2005. Genome-wide prediction and identification of *cis*-natural antisense transcripts in *Arabidopsis thaliana*. *Genome Biol.* **6:** R30.

Wassenegger M., Heimes S., Riedel L., and Sanger H.L. 1994. RNA-directed de novo methylation of genomic sequences in plants. *Cell* **76:** 567–576.

Xie Z., Johansen L.K., Gustafson A.M., Kasschau K.D., Lellis A.D., Zilberman D., Jacobsen S.E., and Carrington J.C. 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2:** E104.

Yoo B.-C., Kragler F., Varkonyi-Gasic E., Haywood V., Archer-Evans S., Lee Y.M., Lough T.J., and Lucas W.J. 2004. A systemic small RNA signalling system in plants. *Plant Cell* **16:** 1979–2000.

## WWW Resources

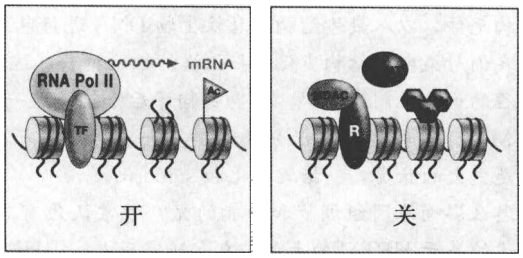
http://asrp.cgrb.oregonstate.edu. *Arabidopsis thaliana* small RNA project

http://mpss.dbi.udel.edu. MPSS (Massively parallel signature sequencing)

http://www.arabidopsis.org/abrc. *Arabidopsis* Biological Resource Center Stocks

http://www.chromdb.org. Plant Chromatin Database

# 第 10 章




# 染色质修饰及其作用机理

Tony Kouzarides[1] and Shelley L. Berger[2]

[1]*The Gurdon Institute, University of Cambridge, Cambridge, United Kingdom*
[2]*The Wistar Institute, Philadelphia, Pennsylvania 19104*

## 本章目录

1. 组蛋白和乙酰化调控转录
2. 乙酰化和去乙酰化
3. 磷酸化
4. 甲基化
   4.1 赖氨酸甲基化
   4.2 赖氨酸去甲基化
   4.3 精氨酸甲基化
5. 瓜氨酸化
6. 泛素化/去泛素化及 Sumo 化
7. 修饰中的旋律
   7.1 组蛋白密码
   7.2 修饰模式
   7.3 与转录激活和延伸相关的染色质结构变化
8. 结束语

参考文献

## 概要

组蛋白是核小体的结构元件，以八聚体结构对真核生物 DNA 进行包装形成染色质。染色质并非均一性结构。近些年，我们对染色质结构变化的知识爆炸性地增长。这进一步增进了我们对以基因组为模板的各种生命进程的调节机制的理解，组蛋白翻译后修饰是这一基因组调节的核心机制。组蛋白翻译后修饰（HPTM）有很多种，可以分成两类。一类是小分子化学基团，包括乙酰化、磷酸化和甲基化。第二类是更大的肽类，包括泛素化和 sumo 化。

组蛋白翻译后修饰是怎么影响基因组调节和功能的呢？通常认为有三种机理，了解这些机理有助于阅读本章中将要论述的关于 HPTM 的大量信息和研究历史。①HPTM 通过某些方式影响染色质的结构，例如，阻碍导致某些染色质构象或高级结构形成的关键相互作用（顺式修饰效应）。与之对应的另两种机理则是反式的。②HPTM 可以破坏染色质或组蛋白结合蛋白的结合。③HPTM也可以提供不同的结合表面给某些效应蛋白。第三种机理有着最详尽的论述。这些招募来的蛋白质决定这最终的功能，即激活或抑制转录。大量已发现的 HPTM 及其不同组合引发了 HPTM 通过不同组合模式进行调节的想法。这些组合模式具有时间顺序，并可在近距离或远距离建立。这些不同的机理决定其不同的功能——某些是瞬时性的，某些则具稳定性并能被表观遗传继承。

20 世纪 60 年代 Vincent Allfrey 鉴定了多种真核生物来源的组蛋白的乙酰化、甲基化和磷酸化。组蛋白还是最早被发现的泛素化的蛋白质底物。然而，尽管 Allfrey 观察到了某些修饰和转录活跃来源的关系，翻译后修饰在基因调控中起作用的遗传学和功能证据很久以后才出现。事实上，在 80 和 90 年代，很多对基因调节进行遗传学和生物化学研究的科学家对组蛋白翻译后修饰在基因调控中哪怕最微小的作用都表示怀疑。

## 1. 组蛋白和乙酰化调控转录

本章将展示，组蛋白是多种翻译后修饰的底物。毫无疑问，会有更多类型的组蛋白翻译后修饰被发现。已知的修饰可以分为本章 2～4 节讨论的小化学基团和第 5 节的更大的组蛋白肽段变化（表 10-1）。图 10-1 中的三种概念模型描述了 HPTM 对染色质模板和相关进程，如基因转录或抑制的调控机理（也见第 3 章第 5 节）。

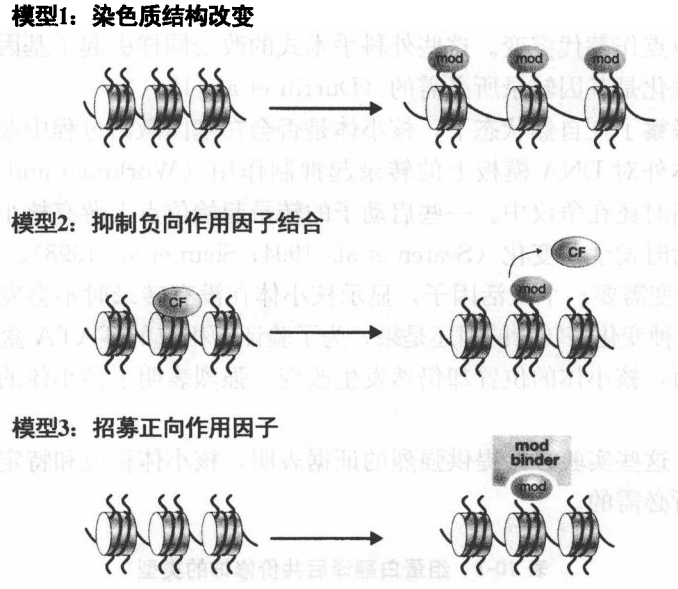



图 10-1　组蛋白翻译后修饰如何影响染色质模板的模型

模型 1 提出染色质结构的改变是由组蛋白乙酰化或磷酸化等顺式共价修饰所介导。模型 2 描述了某种组蛋白翻译后修饰对一个染色质结合因子的抑制作用，以 H3S10 为例子，其磷酸化可拮抗 HP1 与甲基化H3K9 的结合。模型 3 中，一个组蛋白翻译后修饰可以为一个染色质结合因子提供结合特异性。一个经典的例子是 HP1 通过其 chromo 结构域对甲基化的 H3K9 相结合。

从历史角度来看，是什么改变了染色质是用于包装 DNA 的惰性填充材料的主流观点？早期 HPTM 调控转录激活和抑制的证据来自 20 世纪 80 年代末酿酒酵母实验。酿酒酵母为通过遗传学实验（正向和反向遗传学）证明组蛋白的重要性提供了有效的平台。原因在于，不同于其他更高级的真核生物具有的多拷贝组蛋白基因，酵母的单拷贝组蛋白基因具有良好的遗传学操作性。例如，在所有基因组来源组蛋白基因都被敲除的背景中，可以通过携带有筛选标记（如 *URA*3）的附加体（episome）引入某个组蛋白基因的拷贝。用同样的方法可以引入带有其他筛选标记的第二份组蛋白基因。第二份基因拷贝可以经过定点突变，然后，第一个野生型的拷贝可以通过 5-FOA(5-fluoroorotic acid）的处理被选择性丢失（5-FOA 使 *URA3* 的基因产物对细胞有毒害作用）。最终的结果是只有含突变的第二份拷贝基因存在于细胞中，其中可以带有任何有兴趣研究的突变。

在酿酒酵母中，组蛋白基因按照 H3/H4 和 H2A/H2B 成对排列，它们的转录受高度调控以便和 S 期保持协调。每个核小体由一个 H3/H4 四聚体和两个 H2A/H2B 二聚体组装成；当二者之一转录过低或过高时，核小体出现不足。通过遗传学手段对组蛋白表达量进行干预提供了染色质结构对调控表达有重要作用的最初证据。其中一项研究通过正向遗传学筛选出了导致一个标记基因激活的随机突变（Clark-Adams et al. 1988）。这些突变被发现是调节组蛋白相对含量的。另一项研究采用了反向遗传学手段，通过抑制组蛋白基因的表达确切证明了组蛋白对基因转录的调控（Han and Grunstein 1988）。接下来的研究则对富集了组蛋白翻译后修饰位点的组蛋白氨基端尾部进行删除，或者对组蛋白乙酰化位点作替代突变。这些外科手术式的改变同样引起了基因激活程度的急剧降低，暗示乙酰化是基因转录所必需的（Durrin et al. 1991）。

其他研究考察了在自然状态下，核小体是否会在基因激活过程中改变。生化实验显示了核小体在体外对 DNA 模板上的转录起抑制作用（Workman and Roeder，1987），但是体内情况当时还在争议中。一些启动子的转录起始位点上游有核小体存在，这些核小体在基因激活时发生了变化（Svaren et al. 1994；Shim et al. 1998）。以 *PHO*5 基因为例，核小体的改变需要一个激活因子，显示核小体在没有转录时不会发生改变。尽管如此，并不清楚这种变化是转录的因还是果。为了验证，相应的 TATA 盒被删除使得转录无法进行。然而，核小体的位置却仍然发生改变，强烈表明了核小体的改变发生于转录之前。

综上所述，这些实验开始提供强烈的证据表明，核小体移位和特定残基的乙酰化可能是转录激活所必需的。

**表 10-1 组蛋白翻译后共价修饰的类型***

| | 转录中的作用 | 组蛋白修饰位点 |
|---|---|---|
| 组 1 | | |
| 乙酰化 | 激活 | H3(K9, K14, K18, K56) |
| | | H4(K5, K8, K12, K16) |
| | | H2A |
| | | H2B(K6, K7, K16, K17) |
| 磷酸化 | 激活 | H3(S10) |
| 甲基化 | 激活 | H3(K4, K36, K79) |
| | 抑制 | H3(K9, K27) |
| | | H4(K20) |
| 组 2 | | |
| 泛素化 | 激活 | H2B(K123) |
| | 抑制 | H2A(K119) |
| Sumo 化 | 抑制 | H3 (?) |
| | | H4(K5, K8, K12, K16) |
| | | H2A(K126) |
| | | H2B(K6, K7, K16, K17) |

* HPTM 分为两组：第一组列出小分子基团修饰；第二组包含更大的化学修饰。

## 2. 乙酰化和去乙酰化

更多的实验途径进一步提供了乙酰化（与未乙酰化比较）和转录相关的证据。转录活跃或准备转录的染色质区域倾向于“开放”构象，可被核酸酶和微球菌核酸酶分解。这种酶可以消化经提取的完整细胞核 DNA。20 世纪 90 年代早期，研究人员开始使用染色质免疫共沉淀（ChIP）这一有力的技术来分析体内结合于特定 DNA 序列的蛋白质。这一技术使用细胞穿透性化学药物（如甲醛）对结合于 DNA 的蛋白质进行共价交联，然后用超声破碎 DNA：蛋白质复合物为更小的片段。后将感兴趣的 DNA：蛋白质复合物以特定抗体为探针免疫共沉淀。随即进行去交联反应，将与抗体识别蛋白相结合的 DNA 分离出来，再通过放射性标记的 DNA 探针或 PCR 进行鉴定。一个研究组使用这种方法研究鸡红细胞中活跃球蛋白基因处核酸酶高敏位点和组蛋白高乙酰化位点间的相关性，结果证明两者的相关性惊人的高（Hebbes et al. 1994）。在酿酒酵母中，类似的方法被用于研究基因组转录沉默区域，结果发现这些区域只有非常低水平的乙酰化（Braunstein et al. 1993）。相应地，通过遗传操作对基因沉默区进行破坏则导致乙酰化水平上升。

所有这些实验都逐步揭示组蛋白，特别是其可逆的乙酰化位点在基因调控中起一定作用。尽管如此，直到 20 世纪 90 年代中期，第一批细胞核组蛋白乙酰化酶和去乙酰化酶才被鉴定出来。这些研究提供了最确凿和直接的证据表明了这些酶在转录中的作用。第一个组蛋白乙酰化酶（HAT）是从四膜虫转录活跃的大核中分离出来的（大核指巨大且转录活跃的细胞核，与减数分裂的小核相对应）（Brownell et al. 1996）。最关键的方法是检测 HAT 活性的“胶内”实验：来自细胞提取物的总蛋白质在含有组蛋白的 SDS 胶上被分开，这些多肽再经过复性处理，具有 HAT 活性的蛋白质因为催化了放射性标记的乙酰辅酶 A 与固定的组蛋白的反应而在胶上被标记出来。这使得进一步生化分离和纯化该多肽成为可能。鉴定出来的 HAT 与之前在酿酒酵母中分离的与转录激活因子作用的转录共激活因子 Gcn5 同源。与此同时，通过生化纯化分离得到了第一个组蛋白去乙酰化酶（HDAC）（Taunton et al. 1996）。在这个例子中，通过将和酶催化位点结合的抑制剂与固相基质相连，把酶从细胞提取物中纯化出来。这个酶和以前分离到的一个与基因抑制有关的辅助因子基因同源。这些组蛋白乙酰基团代谢酶的发现带来了今天的基因特异性组蛋白翻译后修饰模型的范例：与 DNA 结合的激活因子募集组蛋白乙酰化酶乙酰化核小体组蛋白，而抑制因子募集组蛋白去乙酰化酶去组蛋白乙酰化。这些变化带来了核小体的调整和相应基因表达的上调或下调（图 10-2）。

许多其他著名的共激活子和共抑制子要么含有 HAT 或 HDAC 活性，要么与这类酶结合（Sterner and Berger 2000；Roth et al. 2001）。更重要的是，其 HAT 和 HDAC 的酶活性对其在基因激活和抑制中的作用至关重要。这些酶通常是大的蛋白质复合物的结构和功能组成模块；组蛋白修饰酶活性只是一种功能，其他的功能则包括例如募集 TATA 序列结合蛋白（TBP）（Grant et al. 1998）。有趣的是，某些核内的激素受体既可以作为与 DNA 结合的转录抑制子（当不与激素配体结合时），也能作为转录激活子

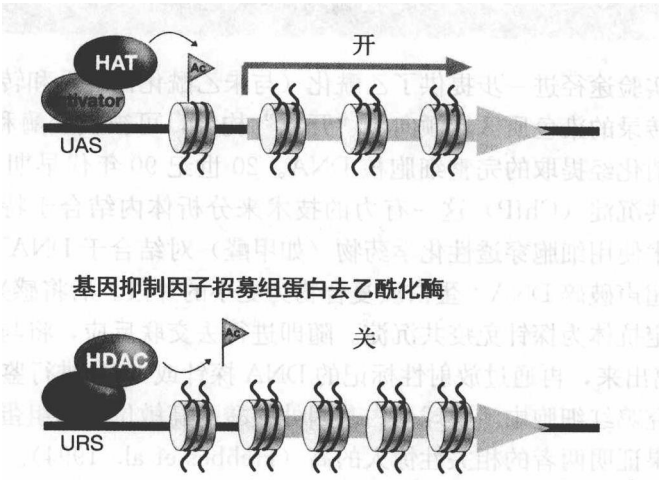



图 10-2 组蛋白修饰酶由 DNA 结合的转录因子招募到启动子

组蛋白乙酰化酶（HAT）结合于特定上游激活序列（UAS）并招募激活因子。它催化附近的组蛋白乙酰化，促进基因激活。组蛋白去乙酰化酶（HDAC）结合于上游抑制序列（URS）并招募转录抑制子，它催化附近的组蛋白去乙酰化，抑制转录。

（当与激素配体结合时）；这些受体调节转录的功能有部分是通过目标染色质区域的翻译后修饰实现的，在未与配体结合时募集 HDAC 而与配体结合时则募集 HAT（Baek and Rosenfeld 2004）。

组蛋白乙酰化酶可以乙酰化四种核心组蛋白，不同的酶具有不同的底物特异性（图 10-3；表 10-1），虽然几乎没有一个酶对应于单一位点。一个主要的组蛋白乙酰化酶家族——GNAT（代表 Gcn5 相关的乙酰化酶）——以组蛋白 H3 作为主要底物。第二个主要家族是 MYST 家族，主要底物是组蛋白 H4。第三个主要家族是以 H3 和 H4 为底物的 CBP/p300，是最复杂的一族。前两个家族（GNAT 和 MYST）的催化结构域的结构分析显示它们是不同的；CBP/p300 家族的结构还没有解开。顺便提一下，每个乙酰化酶家族都具有乙酰化非组蛋白底物的能力（Glozak et al. 2005）。

如前所述，有三种模型解释了 HPTM 对染色质结构的调节（图 10-1）。第一个模型考虑了 HPTM 直接导致的染色质变化，如电荷的改变。在这种情况中，乙酰化中和了带正电的赖氨酸，降低了强碱性的组蛋白或组蛋白尾巴与带负电的 DNA 的结合强度，并因此开放了 DNA 结合位点（Vettese-Dadey et al. 1996）。有证据支持乙酰化使核小体排列松散，这与其打开染色质、激活基因表达的角色相一致（Shogren-Knaak et al. 2006）。第三个模型提出 HPTM 提供结合位点给某些染色质结合蛋白，调节以 DNA 为模板的一些过程；这首先在乙酰化中得到了印证。常见于染色质结合蛋白中的一个特殊结构域称为 bromo 结构域，特异性地结合乙酰化的赖氨酸（图 10-3）（Dhalluin et al. 1999）。bromo 结构域存在于许多组蛋白乙酰化酶中，如 Gcn5 和 CBP/p300。带有这个模块的蛋白，作为大的染色质结合或调节复合物的一部分时（如 ATP

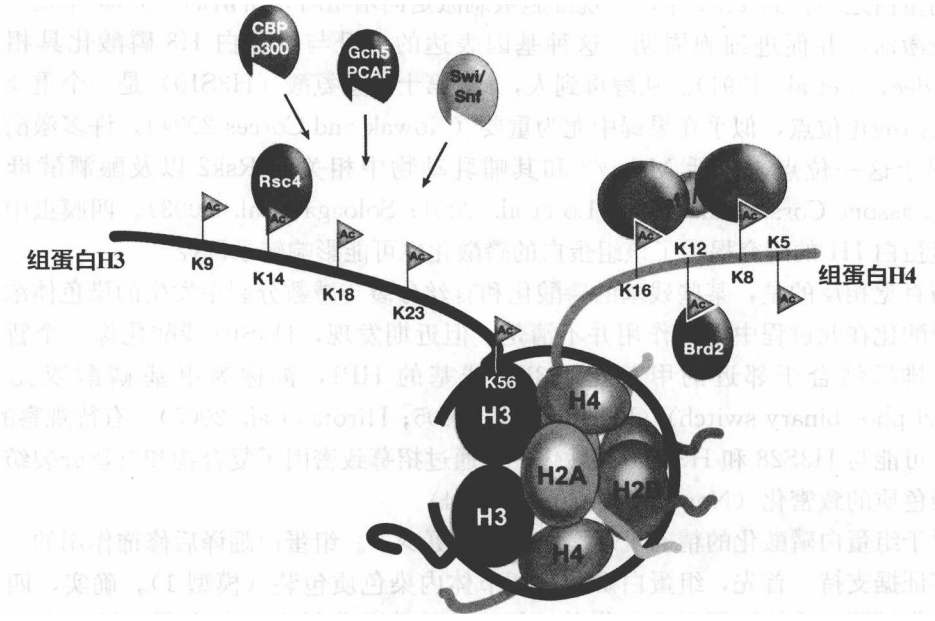



图 10-3　已知组蛋白乙酰化位点

组蛋白乙酰化多发生于 H3 和 H4 氨基端的赖氨酸残基，位于组蛋白球形结构域的 H3K56 是一个例外。具有乙酰化组蛋白结合特异性的蛋白如图示。

依赖的重塑复合物 Swi/Snf）促进其与染色质的结合（Hassan et al. 2002）。其他具有 bromo 结构域并特异结合乙酰化组蛋白的例子有 TFIID 复合物中的 Taf1 和 Bdf1，Rsc 重塑复合物中的 Rsc4 以及 bromo 结构域大家族中的 Brd2。

去乙酰化酶有很多种（Kurdistani and Grunstein 2003；Yang and Seto 2003）。它们分为进化上从酿酒酵母到哺乳动物保守的三种催化类型，称为 I 类、II 类和 III 类（或 Sir2 相关酶）。I 类和 II 类相似都不需要辅酶完成去乙酰化，而 Sir2 相关酶需要 NAD 作为辅酶。这三个家族代表性蛋白的结构都已经被解析。许多 HDAC 存在于大的多亚基复合物中，内含的其他亚基负责将酶定位到目的基因，抑制其转录。例如，Rpd3 是 HDAC Sin3 的大复合物的组分，后者和 DNA 结合的抑制子相作用（Kurdistani and Grunstein 2003；Yang and Seto 2003）。Rpd3 同时又是一个小复合物的组分，这个复合物通过 chromo 结构域结合 H3K36me3，并被定位到基因的开放阅读框（见第 4 章中关于 chromo 结构域的论述）。这导致组蛋白去乙酰化，部分是为抑制基因内 RNA 聚合酶 II 的起始，同时也调节转录周期中的不同阶段（Carrozza et al. 2005；Joshi and Struhl 2005）。

### 3. 磷　酸　化

磷酸化是最著名的翻译后修饰，因为早已知道激酶调节的信号转导可以起始于细胞表面，通过细胞质进入细胞核，从而导致基因表达的变化。组蛋白是最早被发现具有磷

酸化的蛋白之一。到 1991 年，发现细胞被刺激走向增殖时，所谓的“立即-早期”基因被转录激活，并促进细胞周期。这种基因表达的上升与组蛋白 H3 磷酸化具相关性（Mahadeevan et al. 1991）。从酵母到人，H3 第十位丝氨酸（H3S10）是一个重要的转录相关磷酸化位点，似乎在果蝇中尤为重要（Nowak and Corces 2004）。许多激酶被发现作用于这一位点，包括 Msk1/2 和其哺乳动物中相关的 Rsk2 以及酿酒酵母中的 SNF1(Sassone-Corsi et al. 1999；Lo et al. 2001；Soloaga et al. 2003）。四膜虫中关于连接组蛋白 H1 的研究揭示了该组蛋白的磷酸化也可能影响转录调控。

与直觉相反的是，某些残基的磷酸化和有丝分裂与减数分裂中发生的染色体浓缩相关。磷酸化在此过程中如何作用并不清楚，但近期发现，H3S10 磷酸化像一个暂时的开关，排斥结合于邻近的甲基化 H3K9 残基的 HP1，被称为甲基-磷酸双元开关（methyl-phos binary switch）（Fischle et al. 2005；Hirota et al. 2005）。有待观察的是，这是否可能与 H3S28 和 H3T11 磷酸化一起通过招募致密因子复合物和有丝分裂纺锤体影响染色质的致密化（Nowak and Corces 2004）。

对于组蛋白磷酸化的精确机理作用所知就更少了。组蛋白翻译后修饰作用的三个模型均有证据支持。首先，组蛋白磷酸化调节体内染色质包装（模型 1）。确实，四膜虫中的工作证明，连接组蛋白 H1 若干相邻残基的磷酸化导致的负电群区域，影响其与 DNA 结合的亲和性，增加周边染色质环境的转录潜力（Dou and Gorovsky 2002）。支持模型 2 的是，与 DNA 结合的蛋白可以被磷酸化而移走。这最近通过有丝分裂特异性的 H3S10 磷酸化引起的 HP1 结合的亲和性降低得到证明（Fischle et al. 2005；Hirota et al. 2005）。模型 3 的证据来自 14-3-3 接头蛋白，一个磷酸基团结合蛋白，识别可诱导基因启动子处的 H3S10 磷酸化（Macdonald et al. 2005）。

## 4. 甲 基 化

作为一种组蛋白共价修饰，甲基化比任何其他的修饰都更为复杂，因为它可以发生在赖氨酸或精氨酸上。此外，不像表 10-1 组 1 中的任何其他修饰，其对转录的影响可以是正向，也可以是反向的，这取决于甲基化的位点。更深程度的复杂性来自每个残基甲基化状态的多样性。赖氨酸可以被单（me1）、双（me2）或三（me3）甲基化，而精氨酸能被单（me1）或双（me2）甲基化。考虑到 H3、H4、H2A 和 H2B 上共有 24 个已知的甲基化位点，不同的核小体甲基化状态数目是惊人的。这样的甲基化核小体的组合潜能，至少在部分意义上，可能调节复杂和动态的过程，例如转录这样需要顺序和时间精确性的事件（Jenuwein and Allis 2001；Y. Zhang and Reinberg 2001；Lee et al. 2005；Martin and Zhang 2005；Wysocka et al. 2006a）。

### 4.1 赖氨酸甲基化

组蛋白中赖氨酸可以甲基化已经知道几十年了。但其生物学意义直到最近第一个以组蛋白为底物的赖氨酸甲基化酶的发现才开始浮出水面（Rea et al. 2000）。现在，更多的组蛋白赖氨酸甲基化酶被鉴定出来，它们在组蛋白上的修饰位点也得以阐释（Martin

and Zhang 2005)。除了 Dot1，所有的这些酶都具备 SET 结构域，其中含有催化活性区域并结合辅酶 S-腺苷-L-甲硫氨酸（S-adenosyl-L-methionine)。

众多的已知甲基化位点中，目前为止有六个研究地比较清楚：五个在 H3 上（K4，K9，K27，K36，K79)，一个在 H4 上（K20)。大体上，H3K4、H3K36 和 H3K79 的甲基化与转录激活相关，其余的与抑制相联系（表 10-1)。此外，其中的两个位点——甲基化 H3K79 和 H4K20——与 DNA 修复过程有关。分别识别这六个甲基化位点的特异性结合蛋白已被鉴定（图 10-4)。这些蛋白都含有三种甲基化赖氨酸识别结构域中的一种：chromo、tudor 和 PHD 重复结构域。在下文中，将对这些修饰逐一进行详细地讨论。

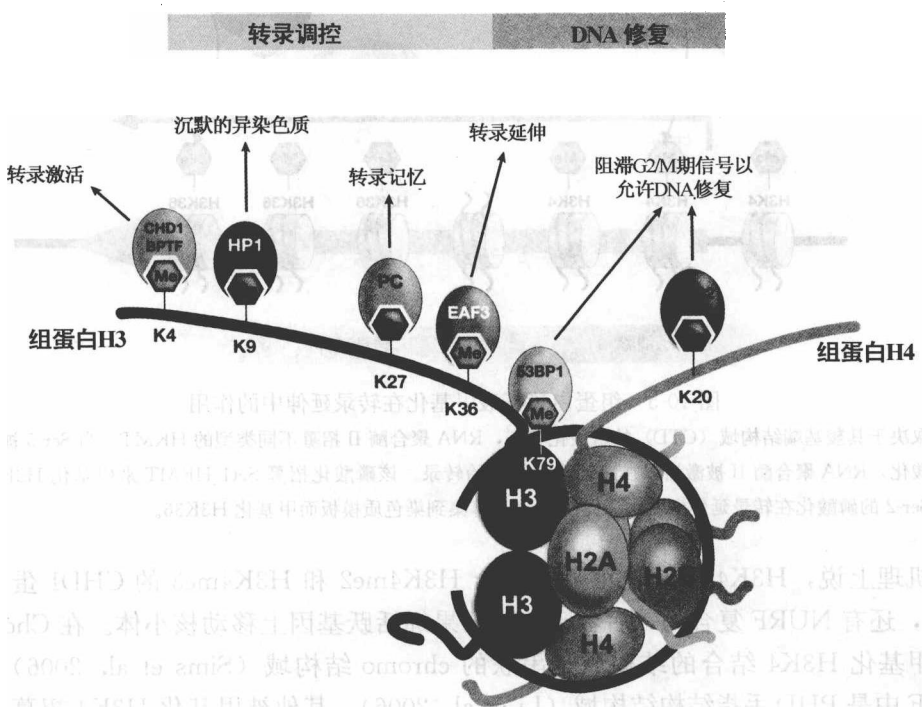



图 10-4　组蛋白甲基化位点，它们的结合蛋白和基因组相关过程中的功能角色

组蛋白甲基化发生在 H3 和 H4 的赖氨酸残基。某些赖氨酸残基和转录激活有关，而其他的与抑制过程相关。结合特定赖氨酸残基的蛋白如图示。

### H3K4 甲基化

H3K4 甲基化富集于常染色质，尤其是活跃转录的或将要激活的基因。H3K4 甲基化和活跃染色质相关的证明来自对鸡 β-球蛋白基因座和酿酒酵母接合型基因座的分析（Litt et al. 2001；Noma et al. 2001)。用甲基化 H3K4 抗体进行染色质免疫共沉淀（ChIP）显示该类修饰伴随活跃基因呈岛状分布。随后在酵母中的工作确立了不同甲基化状态对基因活性的重要性，三甲基化状态（H3K4me3）出现于活跃的转录过程中（Santos-Rosa et al. 2002)。

酵母转录激活时，观察到 H3K4me3 位于基因的 5′端。转录装置的三个组分被认为与此标记有关。首先，RNA 聚合酶 II 羧基端结构域第 5 位丝氨酸的磷酸化可以招募 Set1 组蛋白甲基化酶（HKTM）以甲基化启动子附近的 H3K4（图 10-5）。该磷酸化通常将 RNA 聚合酶 II 从转录起始复合物状态释放进入早期延伸复合物状态（通常称为启动子清除或逃逸）。第二个募集 H3K4me3 的组分是 PAF 复合物，该复合物调节 RNA 代谢的不同步骤并和 Set1 作用。第三个对建立 H3K4me3 起重要作用的组分是 H2B 第 123 位赖氨酸的单泛素化（H2BK123ub1，或人的 H2BK120ub，详细讨论见第 6 节）；但是，H3K4me3 特异性结合因子的作用刚开始得以展示。

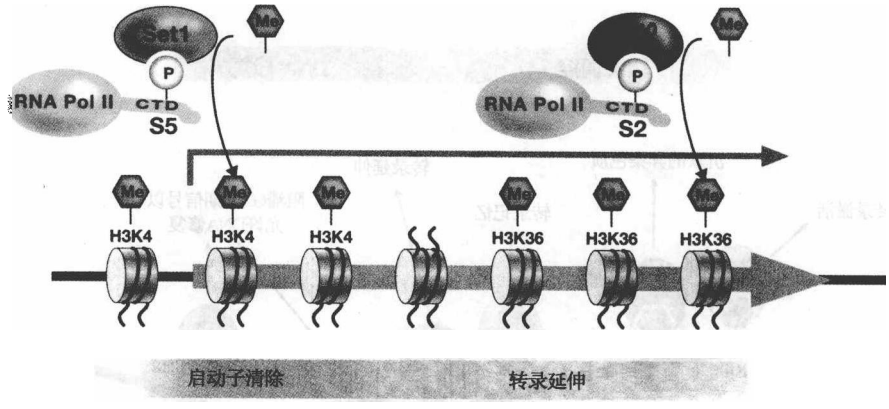



图 10-5 组蛋白赖氨酸甲基化在转录延伸中的作用

取决于其羧基端结构域（CTD）的磷酸化状态，RNA 聚合酶 II 招募不同类型的 HKMT。当 Ser-5 被磷酸化，RNA 聚合酶 II 被激活以在启动子附近起始转录。该磷酸化招募 Set1 HKMT 来甲基化 H3K4。Ser-2 的磷酸化在转录延伸中发生，促使 Set2 被募集到染色质模板而甲基化 H3K36。

机理上说，H3K4 甲基化可募集结合 H3K4me2 和 H3K4me3 的 CHD1 蛋白（图 10-4），还有 NURF 复合物，后者已知能在果蝇活跃基因上移动核小体。在 Chd1 中介导与甲基化 H3K4 结合的结构域是串联的 chromo 结构域（Sims et al. 2006），而在 NURF 中是 PHD 手指结构结构域（Li et al. 2006）。其他被甲基化 H3K4 招募的蛋白还包括通过其他蛋白间接结合的 ISWI ATP 酶。与此相反，有证据表明哺乳动物 NuRD 抑制复合物不再能结合 K4 甲基化的 H3 尾巴（D. Y. Lee et al. 2005；Martin and Zhang 2005）。

H3K4 的甲基化似乎和其他修饰有信息交流。例如，H3K4 被甲基化且 H3S10 同时被磷酸化，则 SUV39H 催化的 H3K9 甲基化在体外无法进行。这可能是一种很好的方式来阻断 H3K9 甲基化对活跃转录基因发生抑制作用。在一种更为精巧的“跨尾反式”交流形式中，H2BK123 的单泛素化影响 H3K4me3 的水平。尚不清楚这是如何发生的，有一种观点是 Set1 复合物不能三甲基化 H3K4 除非核小体处于某种由 H2B 泛素化决定的构象中（Zhang and Reinberg 2001；D. Y. Lee et al. 2005）。

人的 Set1 同源蛋白，Set1/MLL/ALL1/HRX 蛋白，可以被募集到 *HOX* 基因启动子。另一种不同的 H3K4 组蛋白赖氨酸甲基化酶（HKMT）——SMYD3，与转录激活

相关。SMYD3 催化的甲基化还与细胞增殖的诱导相关。事实上，关于人类 H3K4 甲基化酶的有限的分析暗示了它涉及癌症的发生（D. Y. Lee et al. 2005）。

H3K36 甲基化

有证据引发了一个提议，即甲基化的 H3K36 为 RNA 聚合酶 II 有效地延伸通过编码区所必需。与 H3K4 甲基化位于基因 5′端不同，这一修饰富集于活跃基因的编码区。Set2 蛋白是可甲基化 H3K36 的组蛋白赖氨酸甲基化酶。Set2 酶倾向于结合 CTD Ser-2 磷酸化的 RNA 聚合酶 II（图 10-5）。顺便提及，与启动子清除相关的磷酸化状态不同，这种类型的 RNA 聚合酶 II 倾向集中在基因的转录区域和 3′端。这和已有的发现相一致：H3K36me3 分布峰位于活跃转录基因的 3′端。和招募 Set1 一样，Set2 被募集到活跃基因也需要 PAF 复合物的成分。但是，H2B 单泛素化对 H3K36 甲基化有负面的抑制作用（Zhang and Reinberg 2001；Martin and Zhang 2005）。确实，在酵母中被招募到转录基因处的 SAGA 复合物，含有 H2BK123 特异性去泛素化酶 Ubp8。更深入的研究暗示 H2BK123 的泛素化和去泛素化是转录延伸中的一个积极过程。

RNA 聚合酶 II 在编码区的延伸需要核小体的乙酰化。转录调控同样需要抑制隐蔽于编码区中的内藏起始位点发生不当转录起始。为了抑制这个过程，Set2 催化的 H3K36 甲基化为 EAF3 蛋白的 chromo 结构域提供了识别位点，EAF3 继续介导 Rpd3S 组蛋白去乙酰化酶（HDAC）复合物的募集。Rpd3S 的去乙酰化酶活性除去和延伸相关的组蛋白乙酰化，从而抑制内部起始（Carrozza et al. 2005；Joshi and Struhl 2005；Keogh et al. 2005）。甲基化的 H3K36 也被发现低水平地分布在诱导型基因的启动子，但这时它的作用似乎是抑制性的（Zhang and Reinberg 2001）。

H3K79 甲基化

H3K79 的甲基化颇不寻常，因为该修饰位于核小体的核心，而不和大多数其他已描述的修饰一样分布于尾端。全基因组范围的分析显示 H3K79 在酵母常染色质区域被甲基化，并且主要分布在活跃基因的编码区。在更高等真核生物中有限的分析也显示了同样的概况。

哺乳动物中甲基化 H3K79 的酶 hDOT1L 被发现介导 MLL-AF10 融合蛋白引起的白血病发生。尽管如此，到目前为止没有发现蛋白结合 H3K79me，并将其与转录事件相联系。H3K79 甲基化如何在转录激活中作用的唯一机理性证据来自酿酒酵母中的工作。这项工作显示了该修饰以某种方式限制常染色质处的抑制性蛋白例如 Sir2 和 Sir3，从而通过增加其在抑制性染色质区域的浓度起到调节和维持沉默的异染色质的作用。酵母中归于 H3K79 HKMT Dot1 的一种独特功能是作为 DNA 修复检验点。与这项发现一致的是，从人类细胞中鉴定出的蛋白 P53BP1 可以结合甲基化的 H3K79，且在 DNA 修复检查点功能中发生作用（Martin and Zhang 2005）。

H3K9 甲基化

这是目前为止研究最多的组蛋白修饰，主要是因为催化 H3K9 甲基化的酶——

SUV39H1——是第一个被鉴定的组蛋白赖氨酸甲基化酶（Rea et al. 2000）。其果蝇中的同源蛋白，Su(var)3-9 最初被鉴定为位置效应花斑的一个抑制因子，暗示其参与了位置效应花斑现象的沉默机理，后者涉及了异染色质蔓延到相邻常染色质基因的过程（详见第 5 章）。SUV39H1 和植物中以 Rubisco 为底物的甲基化酶序列具有相似性，这一认识导致了将 Suv39 SET 结构域鉴定为 H3K9 甲基化催化结构域。

明确 H3K9 甲基化在着丝粒周围异染色质形成中的功能已经取得进展，这在其他章节中也多有讨论（果蝇见第 5 章；裂殖酵母见第 6 章；关于 RNAi 介导的异染色质形成见第 8 章）。结果主要来自裂殖酵母和哺乳动物中的研究，这两者的异染色质结构被认为有相当的保守性（但是注意，酿酒酵母中并没有发现 H3K9 甲基化）。总的来说，我们认识的第一阶段来自对异染色质建立相关因子的研究。这包括两个蛋白的协作：SUV39H（或裂殖酵母中的 Clr4）及其结合物 HP1（或裂殖酵母中的 Swi6）（Nakayama et al. 2001；Noma et al. 2001）。SUV39H 甲基化 H3K9，为 HP1 提供通过 chromo 结构域作用的结合平台的模型已被提出（Bannister et al. 2001；Lachner et al. 2001）。一旦 HP1 结合，它可以通过与 SUV39H 的相互作用进一步催化相邻的组蛋白甲基化，向邻近的核小体蔓延（Nakayama et al. 2001）。此外，HP1 相互间通过 chromo shadow 结构域的结合也促进异染色质的蔓延。但是，HP1 蔓延是如何控制致密包装的异染色质结构的形成还不清楚。

上述模型预言了应该存在一种异染色质特异性的 SUV39H1 组蛋白赖氨酸甲基化酶招募机制，且这一机制应发生于 HP1 蔓延之前。裂殖酵母中一系列实验表明异染色质形成和小干扰 RNA（siRNA）的生成有关，为解释这一机制提供了线索（Hall et al. 2002；Volpe et al. 2002）。这些 RNA 由着丝粒区域重复序列双向转录后被切割酶 dicer 加工而成。随后 siRNA 被包装进入 RITS 复合物，其中含有结合甲基化 H3K9 的 chromo 结构域蛋白，chp1。因此，RITS 复合物定位到染色质成为异染色质形成的起始阶段。上文所述的，在 20 kb 区域内异染色质的蔓延和维持需要组蛋白甲基化酶 Clr4 介导的组蛋白 H3K9 甲基化以及 Swi6 结合到 H3K9 甲基化的染色质（Martin and Zhang 2005；详见第 8 章）。

不同抑制性表观遗传机制间的互相依赖关系最初在粗糙脉胞菌（*Neurospora Crassa*）研究中被发现，随后在植物研究中也被证实，尤其是证明了 H3K9 甲基化和 DNA 甲基化间的联系（见第 6、8 章）。H3K9 甲基化是 DNA 甲基化的前提，反之亦然，H3K9 甲基化也依赖于 DNA 甲基化。进一步，近期在缺失 DNA 甲基化酶（DNMT）的哺乳动物癌细胞中的研究显示了 H3K9 甲基化水平的降低，而这可以归因于甲基化 CpG 结合蛋白 1(MBD1) 和 H3K9 组蛋白甲基化酶 SETDB1 间的结合（Zhang and Reinberg 2001；Martin and Zhang 2005；详见第 18 章）。

H3K9 甲基化在常染色质基因抑制中也有功能。染色质免疫共沉淀（ChIP）实验在哺乳动物沉默基因的启动子区域检测到该甲基化。这种在常染色质位点的抑制作用的机理似乎稍微有别于在异染色质处的情况。RB 抑制蛋白将组蛋白赖氨酸甲基化酶 SUV39H1 和 HP1 送到常染色质基因处，例如 E2F 调控的周期素 E（cyclinE）基因。与异染色质不同的是，HP1 似乎只占据转录起始位点周围一个或几个核小体，虽然

H3K9 甲基化在启动子其他区域也存在。在另一个例子中，KAP1 抑制因子把 ESET/SETDB1 组蛋白赖氨酸甲基化酶带到受 KAP1 调控的基因启动子处，并通过甲基化 H3K9 和招募 HP1 沉默转录。HP1 在这些常染色质基因启动子分布的特殊限制和蔓延的预防，表明 HP1 的作用机制与在异染色质处不同。一个有某些证据支持的 HP1 作用机制的可能性是，HP1 通过将靶基因定位到细胞核内异染色质丰富的区域来实现其功能。已经观察到常染色质基因在沉默时，被移到异染色质区域的现象，这个过程依赖于 HP1 的 γ 亚型（Martin and Zhang 2005）。

端粒处异染色质形成虽然也涉及 HP1 和 H3K9 甲基化，但与前文所述的着丝粒周边区域和沉默的常染色质区域有着区别。在果蝇中，HP1 并非由其 chromo 或 chromoshadow 结构域被募集到端粒末端，H3K9 甲基化是由一个未知的组蛋白赖氨酸甲基化酶所催化。在哺乳动物中，CBX1、CBX3 和 CBX5 等不同的 HP1 同源物参与结合被 SUV39H1 和 SUV39H2 蛋白产生的甲基化 H3K9，在染色体末端形成抑制性染色质区域（详见第 14 章）。

H3K27 甲基化

H3K27 甲基化是一种抑制性修饰，出现在细胞中三个不同的地方：①常染色质基因座，在果蝇中主要是含有 Polycomb 应答元件（PRE）处；②着丝粒周边异染色质；③哺乳动物中失活的 X 染色体中。

人类细胞中介导 H3K27 甲基化的酶是 EZH2，是果蝇中 *ZESTE* 增强子［E(Z)］的同源蛋白。EZH2 酶存在于许多不同的 Polycomb 抑制性复合物（PRC）中，这些复合物与果蝇启动子中特异的抑制性 Polycomb DNA 元件相结合（又见第 11 章）。含 EZH2 的复合物如何定位到哺乳动物中的特定基因尚不清楚，因为哺乳动物中 Polycomb 抑制性元件还未被鉴定。尽管如此，EZH2 复合物的靶向定位可能由多种转录因子介导，包括 GAGA 和 MYC。EZH2 抑制作用的机理包括 H3K27 甲基化和招募 Polycomb(Pc) 蛋白到该被修饰的位点（图 10-1 中模型 3）。H3K27 甲基化通路的一个重要方面是其与癌症相关。EZH2 H3K27 组蛋白赖氨酸甲基化酶在多种癌症组织中高表达，如乳腺癌和前列腺癌（Martin and Zhang 2005）。

H4K20 甲基化

关于该修饰在转录控制中角色的机理性认识很少。明确已知的是 H4K20me2 与 H4K20me3 分布在着丝粒周边异染色质，而介导该修饰的组蛋白赖氨酸甲基化酶是 SUV4-20H1 及 SUV4-20H2。H3K9 的甲基化似乎是 H4K20 甲基化所必需的。

高等真核生物中另一可单甲基化 H4K20 的酶是 PR-Set7。该酶被认为与有丝分裂相关。最后，有功能性证据支持在裂殖酵母（原文误为酿酒酵母，译者注）中 H4K20 甲基化通过结合 DNA 损伤检验点蛋白 CrB2 与 DNA 修复相关（Martin and Zhang 2005）。

## 4.2　赖氨酸去甲基化

直到最近，我们还不清楚组蛋白赖氨酸去甲基化在细胞中是否发生。对这种酶的搜寻一直无果，同时存在异染色质区域甲基基团相当稳定的证据。LSD1 的发现改变了所有这

些（Shi et al. 2004）。该蛋白是一个特异性去除 H3K4 甲基基团的酶，并参与抑制作用。LSD1 存在于多种不同的抑制性复合物，其中一些有助于 LSD1 更有效地去甲基化核小体内组蛋白 H3 上的赖氨酸（M. G. Lee et al. 2005；Shi et al. 2005）。当 LSD1 与雄激素受体（androgen receptor，AR）结合时，其特异性会发生改变。LSD1-AR 复合物去甲基化 H3K9 而非 H3K4，并且在这种情况下，激活而非抑制转录（Metzger et al. 2005）。

最近，发现了五种新的去甲基化酶，它们共同具有一个有别于 LSD1 的 JmjC 催化结构域（图 10-6）。该结构域曾被预测具有酶活性（Trewick et al. 2005）。这些新的去甲基化酶对不同甲基化状态的 H3K9 和 H3K36 起作用。JHDM1 去甲基化 H3K36me1 和 me2；JHDM2A 去甲基化 H3K9me1 和 me2（Tsukada et al. 2006；Yamane et al. 2006）。这两个被修饰残基的三甲基化状态则由另一套酶去除。JHDM3A 和 JMJD2A 可以同时作用于 H3K36me3 和 H3K9me3（Cloos et al. 2006；Tsukada et al. 2006；Whetstine et al. 2006）。或许令人吃惊的是存在能同时去甲基化一个活跃标记（如 H3K36me3）和一个抑制性标记（如 H3K9me3）的酶。最近关于 H3K9 甲基化也和活跃转录基因有关的发现可能有助于解释这一现象（Vakoc et al. 2005）。在更传统的机理中，JMJD2A 酶被雄激素受体（AR）募集到启动子，通过去甲基化 H3K9 参与激活（Yamane et al. 2006）。JMJD2A 的结构分析显示其四个不同结构域（JmjN，JmjC，一个非经典的锌指和一个羧基端结构域）共同组成催化核心。这些结构域一起形成的深沟需要底物或酶的构象变化以容纳甲基基团进行去甲基化。这种构象变化也许能解释去甲基化的专一性（Chen et al. 2006）。

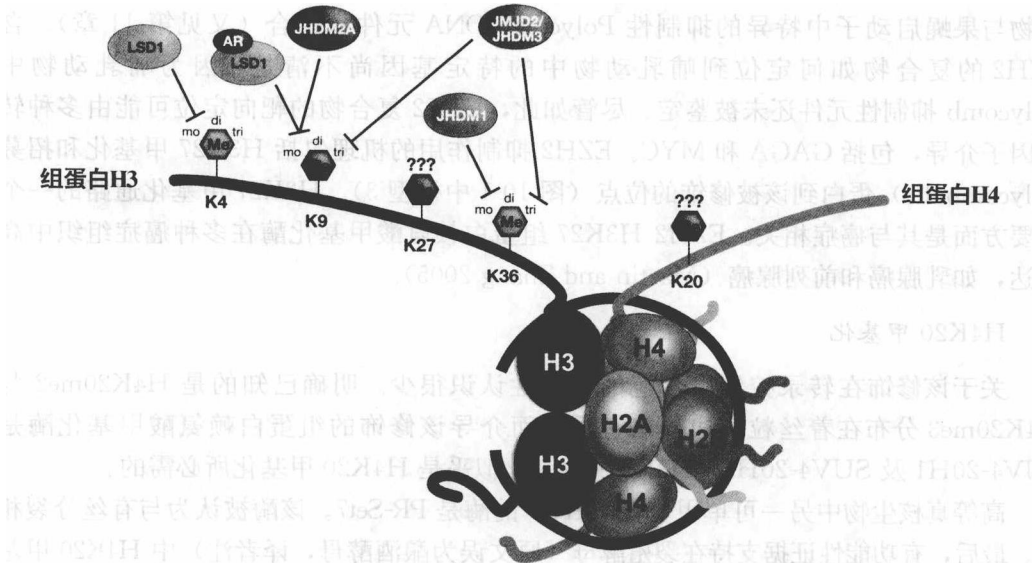



图 10-6 组蛋白赖氨酸去甲基化酶及其在组蛋白 H3 上的去甲基化位点

组蛋白甲基化位点可以是单、双或三甲基化。已知的组蛋白赖氨酸去甲基化酶显示出对组蛋白残基位置及甲基化水平的特异性。

有趣的是新发现的去甲基化酶之一 JMJD2C（以前称为 GASC1），是一个在鳞片状

恶性肿瘤中被扩增的基因。与该酶在癌症发生中的致病作用相一致，GASC1 过表达会诱导细胞增殖（Cloos et al. 2006）。这些结果一起意味着去甲基化酶和组蛋白赖氨酸甲基化酶都可以作为抗癌药物研发的靶标（也见第 24 章）。

### 4.3　精氨酸甲基化

组蛋白精氨酸甲基化在转录调控中的重要性在 CARM1 被鉴定后才被认识到，这是一个在体外甲基化 H3 的精氨酸的酶（Chen et al. 1999）。在体内，精氨酸甲基化在实验中被甲基化精氨酸专一性抗体所证实（Strahl et al. 2001；Wang et al. 2001；Bauer et al. 2002）。

精氨酸甲基化被认为参与转录的正负调控。PRMT1(protein arginine methyltransferase) 和 PRMT4/CARM1 这两个甲基化酶与转录激活相关。PRMT1 可以甲基化 H4R3(Strahl et al. 2001；Wang et al. 2001)，而 PRMT4/CARM1 能催化 H3R2、H3R17 和 H3R26 的甲基化（Schurter et al. 2001；Bauer et al. 2002）。特定的转录因子（NR、p53、YY1、NFKB）招募这些酶到特异的启动子以激活转录。相反，PRMT5（能甲基化 H3R8 和 H4R3）是很多基因的抑制因子，其中包括一些受 *MYC* 调控的基因（Fabbrizio et al. 2002；Pal et al. 2003）。

我们关于精氨酸甲基化的大部分知识来自调控 *pS2* 基因的雌激素信号通路的分析。染色质免疫共沉淀显示该基因被雌激素激活后经历了一系列复杂和周期性的事件。*pS2* 基因接受刺激几分钟内，雌激素受体先被募集到其启动子，同时带来许多修饰组蛋白的蛋白复合物和酶（Metivier et al. 2003）。与此相关的是募集能甲基化组蛋白 H3 和 H4 精氨酸的 CARM1 及 PRMT1（Ma et al. 2001；Bauer et al. 2002）。该甲基化在酶到位后很快就被检测到，并且与启动子上活跃的 RNA 聚合酶 II 的出现相一致。奇怪的是，尽管如此，这些事件后几分钟，抗体不再能检测到精氨酸甲基化，而且 RNA 聚合酶 II 消失了。紧接着，精氨酸甲基化和 RNA 聚合酶 II 又重新出现（Metivier et al. 2003；Cuthbert et al. 2004；Wang et al. 2004）。这些周期性变化事件的原因并不清楚。一种可能性是它提供了一种机制可以在雌激素信号失败的时候快速关闭转录。

重组染色质模板上的实验帮助确立了精氨酸甲基化在基因表达中的直接作用。对 p53 激活性转录的体外分析显示 PRMT1 和 CARM1 介导的甲基化与 CBP/p300 介导的乙酰化具有协同效应（An et al. 2004）。进一步，这些体外的分析确认了体内实验观察到的 pS2 基因激活时的一系列有序事件，其中 PRMT1，CBP/p300 和 CARM1 的顺序活性为必须（Metivier et al. 2003）。

鉴于精氨酸甲基化是如此动态的一个过程，多种控制精氨酸甲基化酶有效性的方式已被描述。首先，酶和另一蛋白的作用可以控制其底物专一性。其次，对于给定的底物，不同的酶之间存在潜在的竞争性。PRMT1 和 PRMT5 都可以甲基化 H4R3，但前者是转录激活因子，后者是抑制因子。甲基化状态的第三层调控可能来自精氨酸去甲基化。这样的活性尚未被分离到，但已有明确的迹象表明精氨酸的甲基基团可以快速消失，使得这种活性的存在成为一个非常诱人的可能性（Zhang and Reinberg 2001；D. Y. Lee et al. 2005；Wysocka et al. 2006a）。

## 5. 瓜氨酸化

精氨酸去甲基化酶的缺少导致了其他类型的酶促反应可能拮抗精氨酸甲基化的提议(Bannister et al. 2002)。其中之一是瓜氨酸化，精氨酸可以脱去一个亚胺转变为瓜氨酸。如果精氨酸是单甲基化的，甲胺基的移除可以有效地脱去精氨酸的甲基基团。组蛋白中瓜氨酸的存在已被证实，而且相关的酶 PADI4 被证明可以将组蛋白内精氨酸转变为瓜氨酸（Cuthbert et al. 2004；Wang et al. 2004)。此外，组蛋白 H3 和 H4 上瓜氨酸的出现伴随着体内精氨酸甲基化的消失。另外，雌激素调控启动子的精氨酸甲基化状态和转录活性正相关，通过对这些启动子的分析表明，瓜氨酸在这些启动子被关闭时出现。和这个修饰有关的很多问题还有待解答。瓜氨酸的作用是抑制精氨酸的甲基化，还是通过主动招募蛋白来抑制转录？瓜氨酸积累又如何被逆转？这显然在启动子上迅速发生，但是由一个酶来催化其逆反应还是仅仅由含有精氨酸的组蛋白变体来对核小体中组蛋白进行置换？

## 6. 泛素化/去泛素化及 Sumo 化

泛素和 SUMO 与翻译后修饰的乙酰化、磷酸化和甲基化差别很大。后三者均为小化学基团翻译后修饰，而 Ub 和 SUMO 则是大的多肽，可以将组蛋白增大约三分之二。Ub 和 SUMO 序列上有 18%相同，且具有同一种三维结构，但其表面电荷并不相同。

组蛋白是最早发现能被单泛素化的蛋白，泛素基团的精确位置最近也得以确定(Robzyk et al. 2000；Wang et al. 2004)。类似于甲基化，而不同于乙酰化和磷酸化（可能还有 sumo 化)，泛素化因特异性位点不同可以起抑制作用或激活作用。与和蛋白降解有关的多泛素化相对，H2A 和 H2B 为单泛素化。各核心组蛋白上的单泛素化效应是相反的(图 10-7)。Rad6/Brel（人类为 RNF20/RNF40+UbcH6）催化的 H2B 单泛素化激活转录(Wood et al. 2003；Kim et al. 2005；Zhu et al. 2005)，并引起 H3K4 甲基化，这在前一节和后一节中均有描述（Henry et al. 2003；Kao et al. 2004)。这一顺序性事件从酵母到人都是保守的，虽然机理尚不清楚（Kim et al. 2005；Zhu et al. 2005)。另一方面，H2AK119ubl 在哺乳动物中由 Polycomb 家族蛋白 Bmil/Ringl A 催化并抑制转录(Wang et al. 2004)。在酵母中未发现抑制性 H2AUb 的进化保守性证据。

到目前为止，还没有鉴定出组蛋白特异的泛素结合蛋白。然而，已知大量的泛素作用结构域可结合非组蛋白泛素化底物，似乎非常可能会找到泛素化组蛋白的效应蛋白。但是，它们可能不同于作用于乙酰化和甲基化的染色质结合结构域；例如，可能同时具有两个结合作用，一个结合泛素表面，另一个作用于组蛋白序列上，从而提供相互作用的特异性。

H2BK123 的去泛素化通过两个不同的蛋白酶（Ubp8 和 Ubp10）起作用，分别参与基因激活和异染色质的维持。Ubp8 是 SAGA 组蛋白乙酰化复合物的一个亚基(Sanders et al. 2002)，在 Rad6 催化泛素化以后发生作用（Henry et al. 2003；Daniel

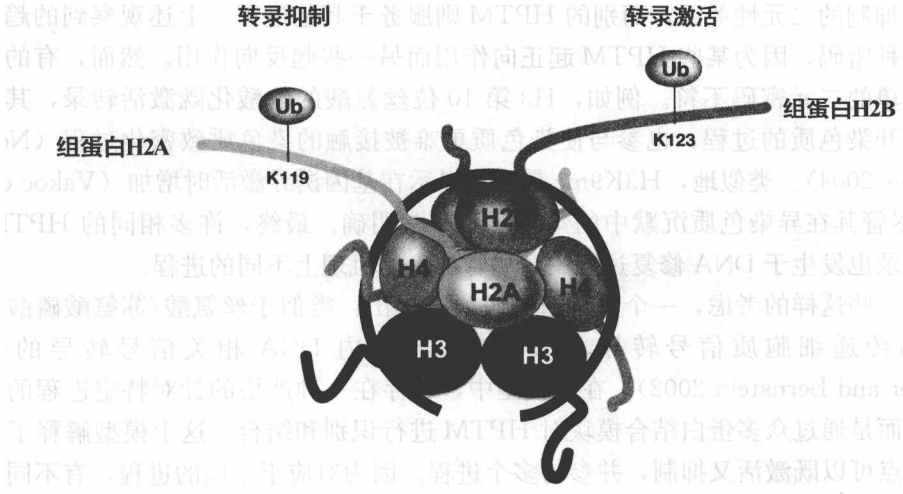



图 10-7　组蛋白泛素化位点及其对转录的调节

H2AK119 的泛素化和转录抑制相关。H2BK123 泛素化则相反，与转录激活有关。

et al. 2004)。H2B 泛素化和去泛素化的顺序发生是建立适当水平的 H3K4（需要 H2Bub）和 H3K36 甲基化（不需要 H2Bub）所必需的（Henry et al. 2003)。Ubp10 作用于沉默区域以维持低水平的 H3K4me 和 H3/H4 赖氨酸乙酰化，从而帮助防止转录（Emre et al. 2005；Gardner et al. 2005)。

Sumo 化是在酵母中发现的唯一一种抑制性组蛋白翻译后修饰，且在哺乳动物中保守（Shiio and Eisenman 2003)。其作用可能通常是负面的——防止激活性组蛋白翻译后修饰（HPTM）的发生。对激活性 HPTM 的抑制可能通过两种机制发生。首先，SUMO 化可以直接封闭组蛋白上的赖氨酸位点，该位点既可能被乙酰化又可能被 SUMO 化（如图 10-1 中的模型 2)。其次，SUMO 化的组蛋白可以通过结合于 DNA 的抑制因子上的 SUMO 基团，为染色质招募组蛋白去乙酰化酶（模型 3)。

## 7. 修饰中的旋律

前文关于转录中众多类型和位点的组蛋白翻译后修饰的讨论可能会引出以下结论，即很少会有能覆盖所有修饰的基本原则。尽管如此，确实存在一些重复响起的主要旋律，尽管其细节可能随组蛋白种类、组蛋白翻译后修饰位点和结合蛋白而改变。的确，染色质调控在启动子之间以及不同通路之间可能存在区别。

### 7.1　组蛋白密码

在对错综复杂的 HPTM 进行了长篇讨论之后，浮现出了一个关键性问题：为什么有如此之多的修饰？很明显它们中的许多和转录相关，其他的则发生在不同的以 DNA 为模板的进程中。因此，一个假说是存在一种组蛋白“密码”，把特定的修饰与各种生

命进程相联系（Strahl and Allis 2000；Turner 2000）。最简单的密码是 HPTM 参与转录激活或抑制的二元性关系，而别的 HPTM 则服务于其他进程。上述观察到的趋势支持这样一种密码，因为某些 HPTM 起正向作用而另一些起反向作用。然而，有的发现与一种简单的二元密码不符。例如，H3 第 10 位丝氨酸的磷酸化既激活转录，其中可能包含打开染色质的过程，也参与使染色质更难被接触的染色质致密化过程（Nowak and Corces 2004）。类似地，H3K9me 最近被显示在基因诱导激活时增加（Vakoc et al. 2005），尽管其在异染色质沉默中的角色已经非常明确。最终，许多相同的 HPTM 既发生于转录也发生于 DNA 修复过程中，尽管两者是机理上不同的进程。

基于一些这样的考虑，一个更普遍的假说被提出：类似于丝氨酸/苏氨酸磷酸化依赖产生和传递细胞质信号转导，HPTM 起着核内 DNA 相关信号转导的作用（Schreiber and Bernstein 2002）。在此模型中，不存在一种严格的针对特定进程的组蛋白密码，而是通过众多蛋白结合模块对 HPTM 进行识别和结合。这个模型解释了为什么任一位点可以既激活又抑制，并参与多个进程。因为对应于不同的进程，有不同的效应蛋白来识别同一个 HPTM。

## 7.2 修饰模式

一些实验证据指出某些 HPTM 引起的染色质结构变化（模型 1）。这可能来自单个或者成群的组蛋白残基电荷的变化。残基被乙酰化或磷酸化时更为如此，因为它们降低组蛋白区域的正电荷（见第 3 章第 5 节）。这样的顺式变化可以改变核小体间距并降低组蛋白和带负电 DNA 的亲和性，例如发生于连接组蛋白上的负电荷区域（Dou and Gorovsky 2002）。这些类型的 HPTM 在影响某些进程，如转录激活或形成更高级的染色质结构时，可能具有累积效应，而非产生二元的开/关效应（Kurdistani et al. 2004；Henikoff 2005）。

无数 HPTM 的另一种“输出”模型认为密码是复杂的，且以模式的形式，按时间顺序得以解读。在这种观点看来，不同模式间在三维空间和时间上错综复杂的关系需要许多染色质相关进程（如转录），来产生机理上有意义的结果。两类 HPTM 模式已被发现。一是在有同一组蛋白尾巴上的模式（顺式模式）；二是在不同组蛋白尾巴上的模式（反式模式）。最清楚的顺式模式是 H3 氨基端的 H3S10ph 和 H3K14ac（Cheung et al. 2000；Lo et al. 2000），其中 H3S10ph 引起 H3K14ac。这种模式建立的机理已在结构细节上得到了解：进行乙酰化的酶和预先磷酸化的 H3 末端有更多的氨基酸侧链接触位点，具有更高的亲和性（Clements et al. 2003）。其他顺式模式有 H3K23 乙酰化、H3R17 甲基化（Daujat et al. 2002）以及 H4R3 甲基化、H4K8 乙酰化（Wang et al. 2001）。

已知有一个反式末端模式，即 H2BK123 泛素化导致 H3K4 甲基化（Briggs et al. 2002；Dover et al. 2002；Sun and Allis 2002）。虽然存在一些可能的假说，联系这两个 HPTM 的机理并不清楚。由于这一联系发生在一个大的修饰（泛素）和一个非相邻的 HPTM 之间，一个模型是泛素像撬棍一样打开染色质，使执行甲基化的酶到达作用位点。第二个模型认为 H2BK123ub1 像其他 HPTM 一样招募效应蛋白。蛋白酶体的非催化部分需要 H2B 泛素化来完成与染色质的结合（Ezhkova and Tansey 2004），同时延伸

复合物 FACT 的功能受 H2B 泛素化刺激（Pavri et al. 2006），虽然这两个复合物与泛素化 H2B 的直接结合有待于证实。

## 7.3　与转录激活和延伸相关的染色质结构变化

转录活跃的常染色质区域含有“未折叠”状态的核小体，称为“珠串”或 11nm 纤维。这种状态的核小体仍然对转录装置有内在的抑制作用。有些转录因子、激活因子或抑制因子能结合到它们核小体上，另一些则不能。此外，结合 DNA 的调控因子招募负责递送 RNA pol II 到启动子的装置受核小体存在的限制。改变染色质结构、调整基因以供转录或促进起始或延伸存在多种机理。部分如图 10-8 所示。

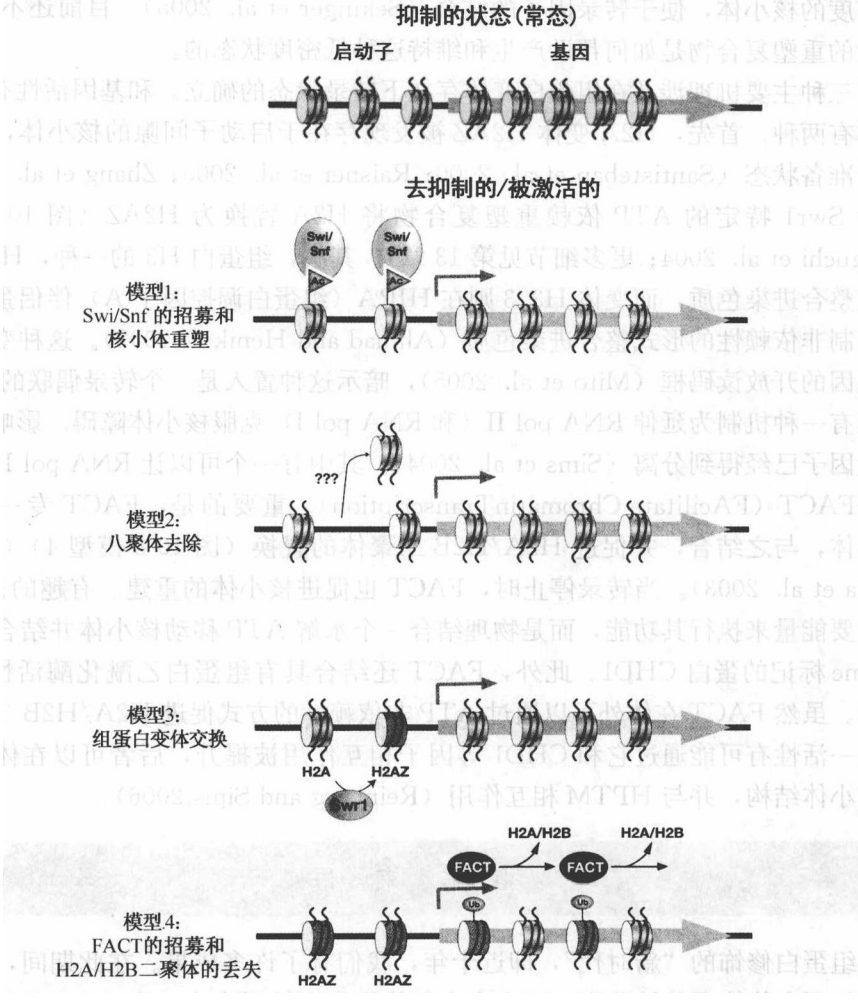



图 10-8　转录过程中染色质重塑和组蛋白交换作用的模型

模型 1 中，Swi/Snf 家族的 ATPase 通过 bromo 结构域识别乙酰化组蛋白并结合染色质，改变所在染色质的结构。模型 2 描述了通过未知机理发生在某些基因座（如 *PHO5*）的八聚体逐出。模型 3 中，ATPase SWR1 催化 H2AZ 置换 H2A，前者使染色质处于转录准备状态。模型 4 集中于转录延伸中 FACT 的作用，其通过置换 H2A/H2B 二聚体帮助核小体拆开。与此同时，组蛋白 H3 在这个过程中可能与 H3. 3 发生交换。

转录中核小体的问题部分通过招募蛋白复合物来移动核小体和（或）改变核小体结构得以解决。这些复合物分为两个家族：一个以 SNF2H（酵母中的 ISW1 和 ISW2）为代表，另一个以 Brahma-Swi/Snf 家族为代表（Narlikar et al. 2002；Peterson 2002；Flaus and Owen-Hughes 2004）。第一个家族移动核小体，而 Swi/Snf 能短暂地改变核小体结构。乙酰化的核小体通过 bromo 结构域作用被 Swi/Snf 复合物所识别（图 10-8 模型 1）。

基因激活中的第二种机理是启动子区域八聚体的选择性丢失。例如，酿酒酵母 *PHO5* 基因受转录诱导时，其启动子区的组蛋白八聚体被去除（图 10-8 模型 2）（Boeger et al. 2003；Reinke and Horz 2003）。另外，酿酒酵母启动子区域组成性地具有低密度的核小体，便于转录因子的结合（Sekinger et al. 2005）。目前还不清楚 ATP 依赖性的重塑复合物是如何帮助产生和维持这种低密度状态的。

第三种主要机理涉及在组蛋白变体存在下转录状态的确立。和基因活性有关的组蛋白变体有两种。首先，H2A 变体 H2AZ 被发现存在于启动子间隙的核小体，使基因处于激活准备状态（Santisteban et al. 2000；Raisner et al. 2005；Zhang et al. 2005）；一个称为 Swr1 特定的 ATP 依赖重塑复合物将 H2A 置换为 H2AZ（图 10-8 模型 3）（Mizuguchi et al. 2004；更多细节见第 13 章）。其次，组蛋白 H3 的一种，H3.1，在复制时被整合进染色质，而变体 H3.3 则在 HIRA（组蛋白调控因子 A）伴侣蛋白的帮助下以复制非依赖性的形式整合进染色质（Ahmad and Henikoff 2002）。这种变体广泛存在于基因的开放读码框（Mito et al. 2005），暗示这种置入是一个转录偶联的进程。

还有一种机制为延伸 RNA pol II（和 RNA pol I）克服核小体障碍。影响转录延伸的大量因子已经得到分离（Sims et al. 2004）。其中有一个可以让 RNA pol II 越过核小体——FACT（FAcilitate ChromatinTranscription）。重要的是，FACT 专一性地作用于核小体，与之结合，并促进 H2A/H2B 二聚体的置换（图 10-8 模型 4）（Belotserkovskaya et al. 2003）。当转录停止时，FACT 也促进核小体的重建。有趣的是，FACT 并不需要能量来执行其功能，而是物理结合一个水解 ATP 移动核小体并结合于激活性 H3K4me 标记的蛋白 CHD1。此外，FACT 还结合具有组蛋白乙酰化酶活性的 NuA4 复合物。虽然 FACT 在体外可以通过 ATP 非依赖性的方式促进 H2A/H2B 二聚体的置换，这一活性有可能通过它和 CHD1 等因子相互作用被提升，后者可以在体内移动或改变核小体结构，并与 HPTM 相互作用（Reinberg and Sims 2006）。

## 8. 结 束 语

在组蛋白修饰的“新时代”，即近十年，我们有了许多成果。在此期间，有六种不同类型组蛋白修饰通路被鉴定，还有众多的修饰位点得到确认。然而这显然仍是认识的开端。机理上，我们知道修饰影响蛋白的结合，但我们仍不确切了解这些蛋白是如何引起染色质结构重建的。我们仍然不知道是否存在一种修饰密码，还是仅仅是一条信号通路中的一部分。另外，我们缺乏对转录之外的，许多与修饰相关的细胞进程的知识。因此，简而言之，我们开始了解到这一系统的复杂性，但是我们距离弄清楚这复杂性还有

很远的路。然而有一件事是明确的：这值得我们去努力探索，因为组蛋白修饰在正常和病理过程中都起着十分重要的作用。

谢精卫　译
袁　文　校
朱　冰　审校

## 参考文献


Ahmad K. and Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9:** 1191–1200.

An W., Kim J., and Roeder R.G. 2004. Ordered cooperative functions of PRMT1, p300, and CARM1 in transcriptional activation by p53. *Cell* **117:** 735–748.

Baek S.H. and Rosenfeld M.G. 2004. Nuclear receptor coregulators: Their modification codes and regulatory mechanism by translocation. *Biochem. Biophys. Res. Commun.* **319:** 707–714.

Bannister A.J., Schneider R., and Kouzarides T. 2002. Histone modification: Dynamic or static? *Cell* **109:** 801–806.

Bannister A.J., Zegerman P., Partridge J.F., Miska E.A., Thomas J.O., Allshire R.C., and Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410:** 120–124.

Bauer U.M., Daujat S., Nielsen S.J., Nightingale K., and Kouzarides T. 2002. Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep.* **3:** 39–44.

Belotserkovskaya R., Oh S., Bondarenko V.A., Orphanides G., Studitsky V.M., and Reinberg D. 2003. FACT facilitates transcription-dependent nucleosome alteration. *Science* **301:** 1090–1093.

Boeger H., Griesenbeck J., Strattan J.S., and Kornberg R.D. 2003. Nucleosomes unfold completely at a transcriptionally active promoter. *Mol. Cell* **11:** 1587–1598.

Braunstein M., Rose A.B., Holmes S.G., Allis C.D., and Broach J.R. 1993. Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev.* **7:** 592–604.

Briggs S.D., Xiao T., Sun Z.W., Caldwell J.A., Shabanowitz J., Hunt D.F., Allis C.D., and Strahl B.D. 2002. Gene silencing: *Trans*-histone regulatory pathway in chromatin. *Nature* **418:** 498.

Brownell J.E., Zhou J., Ranalli T., Kobayashi R., Edmondson D.G., Roth S.Y., and Allis C.D. 1996. *Tetrahymena* histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84:** 843–851.

Carrozza M.J., Li B., Florens L., Suganuma T., Swanson S.K., Lee K.K., Shia W.J., Anderson S., Yates J., Washburn M.P., and Workman J.L. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123:** 581–592.

Chen D., Ma H., Hong H., Koh S.S., Huang S.M., Schurter B.T., Aswad D.W., and Stallcup M.R. 1999. Regulation of transcription by a protein methyltransferase. *Science* **284:** 2174–2177.

Chen Z., Zang J., Whetstine J., Hong X., Davrazou F., Kutateladze T.G., Simpson M., Mao Q., Pan C.H., Dai S., et al. 2006. Structural insights into histone demethylation by JMJD2 family members. *Cell* **125:** 691–702.

Cheung P., Tanner K.G., Cheung W.L., Sassone-Corsi P., Denu J.M., and Allis C.D. 2000. Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol. Cell* **5:** 905–915.

Clark-Adams C.D., Norris D., Osley M.A., Fassler J.S., and Winston F. 1988. Changes in histone gene dosage alter transcription in yeast. *Genes Dev.* **2:** 150–159.

Clements A., Poux A.N., Lo W.S., Pillus L., Berger S.L., and Marmorstein R. 2003. Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase. *Mol. Cell* **12:** 461–473.

Cloos P.A., Christensen J., Agger K., Maiolica A., Rappsilber J., Antal T., Hansen K.H., and Helin K. 2006. The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* **442:** 307–311.

Cuthbert G.L., Daujat S., Snowden A.W., Erdjument-Bromage H., Hagiwara T., Yamada M., Schneider R., Gregory P.D., Tempst P., Bannister A.J., and Kouzarides T. 2004. Histone deimination antagonizes arginine methylation. *Cell* **118:** 545–553.

Daniel J.A., Torok M.S., Sun Z.W., Schieltz D., Allis C.D., Yates J.R., III, and Grant P.A. 2004. Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J. Biol. Chem.* **279:** 1867–1871.

Daujat S., Bauer U.M., Shah V., Turner B., Berger S., and Kouzarides T. 2002. Crosstalk between CARM1 methylation and CBP acetylation on histone H3. *Curr. Biol.* **12:** 2090–2097.

Dhalluin C., Carlson J.E., Zeng L., He C., Aggarwal A.K., and Zhou M.M. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399:** 491–496.

Dou Y. and Gorovsky M.A. 2002. Regulation of transcription by H1 phosphorylation in *Tetrahymena* is position independent and requires clustered sites. *Proc. Natl. Acad. Sci.* **99:** 6142–6146.

Dover J., Schneider J., Tawiah-Boateng M.A., Wood A., Dean K., Johnston M., and Shilatifard A. 2002. Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J. Biol. Chem.* **277:** 28368–28371.

Durrin L.K., Mann R.K., Kayne P.S., and Grunstein M. 1991. Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* **65:** 1023–1031.

Emre N.C., Ingvarsdottir K., Wyce A., Wood A., Krogan N.J., Henry K.W., Li K., Marmorstein R., Greenblatt J.F., Shilatifard A., and Berger S.L. 2005. Maintenance of low histone ubiquitylation by Ubp10 correlates with telomere-proximal Sir2 association and gene silencing. *Mol. Cell* **17:** 585–594.

Ezhkova E. and Tansey W.P. 2004. Proteasomal ATPases link ubiquitylation of histone H2B to methylation of histone H3. *Mol. Cell* **13:** 435–442.

Fabbrizio E., El Messaoudi S., Polanowska J., Paul C., Cook J.R., Lee J.H., Negre V., Rousset M., Pestka S., Le Cam A., and Sardet C. 2002. Negative regulation of transcription by the type II arginine methyltransferase PRMT5. *EMBO Rep.* **3:** 641–645.

Fischle W., Tseng B.S., Dormann H.L., Ueberheide B.M., Garcia B.A., Shabanowitz J., Hunt D.F., Funabiki H., and Allis C.D. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438:** 1116–1122.

Flaus A. and Owen-Hughes T. 2004. Mechanisms for ATP-dependent chromatin remodelling: Farewell to the tuna-can octamer? *Curr. Opin. Genet. Dev.* **14:** 165–173.

Fodor B.D., Kubicek S., Yonezawa M., O'Sullivan R.J., Sengupta R., Perez-Burgos L., Opravil S., Mechtler K., Schotta G., and Jenuwein T. 2006. Jmjd2b antagonizes H3K9 trimethylation at pericentric

heterochromatin in mammalian cells. *Genes Dev.* **20:** 1557–1562.

Gardner R.G., Nelson Z.W., and Gottschling D.E. 2005. Ubp10/Dot4p regulates the persistence of ubiquitinated histone H2B: Distinct roles in telomeric silencing and general chromatin. *Mol. Cell. Biol.* **25:** 6123–6139.

Glozak M.A., Sengupta N., Zhang X., and Seto E. 2005. Acetylation and deacetylation of non-histone proteins. *Gene* **363:** 15–23.

Grant P.A., Sterner D.E., Duggan L.J., Workman J.L., and Berger S.L. 1998. The SAGA unfolds: Convergence of transcription regulators in chromatin-modifying complexes. *Trends Cell Biol.* **8:** 193–197.

Hall I.M., Shankaranarayana G.D., Noma K., Ayoub N., Cohen A., and Grewal S.I. 2002. Establishment and maintenance of a heterochromatin domain. *Science* **297:** 2232–2237.

Hampsey M. and Reinberg D. 2003. Tails of intrigue: Phosphorylation of RNA polymerase II mediates histone methylation. *Cell* **113:** 429–432.

Han M. and Grunstein M. 1988. Nucleosome loss activates yeast downstream promoters in vivo. *Cell* **55:** 1137–1145.

Hassan A.H., Prochasson P., Neely K.E., Galasinski S.C., Chandy M., Carrozza M.J., and Workman J.L. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111:** 369–379.

Hebbes T.R., Clayton A.L., Thorne A.W., and Crane-Robinson C. 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J.* **13:** 1823–1830.

Henikoff S. 2005. Histone modifications: Combinatorial complexity or cumulative simplicity? *Proc. Natl. Acad. Sci.* **102:** 5308–5309.

Henry K.W., Wyce A., Lo W.S., Duggan L.J., Emre N.C., Kao C.F., Pillus L., Shilatifard A., Osley M.A., and Berger S.L. 2003. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev.* **17:** 2648–2663.

Hirota T., Lipp J.J., Toh B.H., and Peters J.M. 2005. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438:** 1176–1180.

Jenuwein T. and Allis C.D. 2001. Translating the histone code. *Science* **293:** 1074–1080.

Joshi A.A. and Struhl K. 2005. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol. Cell* **20:** 971–978.

Kao C.F., Hillyer C., Tsukuda T., Henry K., Berger S., and Osley M.A. 2004. Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B. *Genes Dev.* **18:** 184–195.

Keogh M.C., Kurdistani S.K., Morris S.A., Ahn S.H., Podolny V., Collins S.R., Schuldiner M., Chin K., Punna T., Thompson N.J., et al. 2005. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123:** 593–605.

Kim J., Hake S.B., and Roeder R.G. 2005. The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. *Mol. Cell* **20:** 759–770.

Klose R.J., Yamane K., Bae Y., Zhang D., Erdjument-Bromage H., Tempst P., Wong J., and Zhang Y. 2006. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* **442:** 312–316.

Kurdistani S.K. and Grunstein M. 2003. Histone acetylation and deacetylation in yeast. *Nat. Rev. Mol. Cell. Biol.* **4:** 276–284.

Kurdistani S.K., Tavazoie S., and Grunstein M. 2004. Mapping global histone acetylation patterns to gene expression. *Cell* **117:** 721–733.

Lachner M., O'Carroll D., Rea S., Mechtler K., and Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410:** 116–120.

Lee D.Y., Teyssier C., Strahl B.D., and Stallcup M.R. 2005. Role of protein methylation in regulation of transcription. *Endocr. Rev.* **26:** 147–170.

Lee M.G., Wynder C., Cooch N., and Shiekhattar R. 2005. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* **437:** 432–435.

Li H., Ilin S., Wang W., Duncan E.M., Wysocka J., Allis C.D., and Patel D.J. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442:** 91–95.

Litt M.D., Simpson M., Gaszner M., Allis C.D., and Felsenfeld G. 2001. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* **293:** 2453–2455.

Lo W.S., Duggan L., Emre N.C., Belotserkovskya R., Lane W.S., Shiekhattar R., and Berger S.L. 2001. Snf1—A histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. *Science* **293:** 1142–1146.

Lo W.S., Trievel R.C., Rojas J.R., Duggan L., Hsu J.Y., Allis C.D., Marmorstein R., and Berger S.L. 2000. Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. *Mol. Cell* **5:** 917–926.

Ma H., Baumann C.T., Li H., Strahl B.D., Rice R., Jelinek M.A., Aswad D.W., Allis C.D., Hager G.L., and Stallcup M.R. 2001. Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter. *Curr. Biol.* **11:** 1981–1985.

Macdonald N., Welburn J.P., Noble M.E., Nguyen A., Yaffe M.B., Clynes D., Moggs J.G., Orphanides G., Thomson S., Edmunds J.W., et al. 2005. Molecular basis for the recognition of phosphorylated and phosphoacetylated histone H3 by 14-3-3. *Mol. Cell* **20:** 199–211.

Mahadevan L.C., Willis A.C., and Barratt M.J. 1991. Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. *Cell* **65:** 775–783.

Martin C. and Zhang Y. 2005. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell. Biol.* **6:** 838–849.

Metivier R., Penot G., Hubner M.R., Reid G., Brand H., Kos M., and Gannon F. 2003. Estrogen receptor-α directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. *Cell* **115:** 751–763.

Metzger E., Wissmann M., Yin N., Muller J.M., Schneider R., Peters A.H., Gunther T., Buettner R., and Schule R. 2005. LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature* **437:** 436–439.

Mito Y., Henikoff J.G., and Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* **37:** 1090–1097.

Mizuguchi G., Shen X., Landry J., Wu W.H., Sen S., and Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303:** 343–348.

Nakayama J., Rice J.C., Strahl B.D., Allis C.D., and Grewal S.I. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292:** 110–113.

Narlikar G.J., Fan H.Y., and Kingston R.E. 2002. Cooperation between complexes that regulate chromatin structure and transcription. *Cell* **108:** 475–487.

Noma K., Allis C.D., and Grewal S.I. 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science* **293:** 1150–1155.

Nowak S.J. and Corces V.G. 2004. Phosphorylation of histone H3: A balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* **20:** 214–220.

Pal S., Yun R., Datta A., Lacomis L., Erdjument-Bromage H., Kumar J., Tempst P., and Sif S. 2003. mSin3A/histone deacetylase 2- and PRMT5-containing Brg1 complex is involved in transcriptional repression of the Myc target gene cad. *Mol. Cell. Biol.* **21:** 7475–7487.

Pavri R., Zhu B., Li G., Trojer P., Mandal S., Shilatifard A., and Reinberg D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125:** 703–717.

Peterson C.L. 2002. Chromatin remodeling: Nucleosomes bulging at the seams. *Curr. Biol.* **12:** R245–R247.

Raisner R.M., Hartley P.D., Meneghini M.D., Bao M.Z., Liu C.L., Schreiber S.L., Rando O.J., and Madhani H.D. 2005. Histone variant H2A.Z marks the 5′ ends of both active and inactive genes in euchromatin. *Cell* **123:** 233–248.

Rea S., Eisenhaber F., O'Carroll D., Strahl B.D., Sun Z.W., Schmid M., Opravil S., Mechtler K., Ponting C.P., Allis C.D., and Jenuwein T. 2000. Regulation of chromatin structure by site-specific histone

H3 methyltransferases. *Nature* **406:** 593–599.

Reinberg D. and Sims R.J., III. 2006. de FACTo nucleosome dynamics. *J. Biol. Chem.*(in press).

Reinke H. and Horz W. 2003. Histones are first hyperacetylated and then lose contact with the activated *PHO5* promoter. *Mol. Cell* **11:** 1599–1607.

Robzyk K., Recht J., and Osley M.A. 2000. Rad6-dependent ubiquitination of histone H2B in yeast. *Science* **287:** 501–504.

Roth S.Y., Denu J.M., and Allis C.D. 2001. Histone acetyltransferases. *Annu. Rev. Biochem.* **70:** 81–120.

Sanders S.L., Jennings J., Canutescu A., Link A.J., and Weil P.A. 2002. Proteomics of the eukaryotic transcription machinery: Identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Mol. Cell. Biol.* **22:** 4723–4738.

Santisteban M.S., Kalashnikova T., and Smith M.M. 2000. Histone H2A.Z regulates transcription and is partially redundant with nucleosome remodeling complexes. *Cell* **103:** 411–422.

Santos-Rosa H., Schneider R., Bannister A.J., Sherriff J., Bernstein B.E., Emre N.C., Schreiber S.L., Mellor J., and Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419:** 407–411.

Sassone-Corsi P., Mizzen C.A., Cheung P., Crosio C., Monaco L., Jacquot S., Hanauer A., and Allis C.D. 1999. Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. *Science* **285:** 886–891.

Schreiber S.L. and Bernstein B.E. 2002. Signaling network model of chromatin. *Cell* **111:** 771–778.

Schurter B.T., Koh S.S., Chen D., Bunick G.J., Harp J.M., Hanson B.L., Henschen-Edman A., Mackay D.R., Stallcup M.R., and Aswad D.W. 2001. Methylation of histone H3 by coactivator-associated arginine methyltransferase 1. *Biochemistry* **40:** 5747–5756.

Sekinger E.A., Moqtaderi Z., and Struhl K. 2005. Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell* **18:** 735–748.

Shi Y., Lan F., Matson C., Mulligan P., Whetstine J.R., Cole P.A., Casero R.A., and Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119:** 941–953.

Shi Y.J., Matson C., Lan F., Iwase S., Baba T., and Shi Y. 2005. Regulation of LSD1 histone demethylase activity by its associated factors. *Mol. Cell.* **19:** 857–864.

Shiio Y. and Eisenman R.N. 2003. Histone sumoylation is associated with transcriptional repression. *Proc. Natl. Acad. Sci.* **100:** 13225–13230.

Shim E.Y., Woodcock C., and Zaret K.S. 1998. Nucleosome positioning by the winged helix transcription factor HNF3. *Genes Dev.* **12:** 5–10.

Shogren-Knaak M., Ishii H., Sun J.M., Pazin M.J., Davie J.R., and Peterson C.L. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311:** 844–847.

Sims. R.J., Belotserkovskaya R., and Reinberg D. 2004. Elongation by RNA polymerase II: The short and long of it. *Genes Dev.* **18:** 2437–2468.

Sims R.J., Chen C.F., Santos-Rosa H., Kouzarides T., Patel S.S., and Reinberg D. 2006. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *J. Biol. Chem* **51:** 41789–41792.

Soloaga A., Thomson S., Wiggin G.R., Rampersaud N., Dyson M.H., Hazzalin C.A., Mahadevan L.C., and Arthur J.S. 2003. MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG-14. *EMBO J.* **22:** 2788–2797.

Sterner D.E. and Berger S.L. 2000. Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev.* **64:** 435–459.

Strahl B.D. and Allis C.D. 2000. The language of covalent histone modifications. *Nature* **403:** 41–45.

Strahl B.D., Briggs S.D., Brame C.J., Caldwell J.A., Koh S.S., Ma H., Cook R.G., Shabanowitz J., Hunt D.F., Stallcup M.R., and Allis C.D. 2001. Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol.* **26:** 996–1000.

Sun Z.W. and Allis C.D. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418:** 104–108.

Svaren J., Schmitz J., and Horz W. 1994. The transactivation domain of Pho4 is required for nucleosome disruption at the *PHO5* promoter. *EMBO J.* **13:** 4856–4862.

Taunton J., Hassig C.A., and Schreiber S.L. 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272:** 408–411.

Trewick S.C., McLaughlin P.J., and Allshire R.C. 2005. Methylation: Lost in hydroxylation? *EMBO Rep.* **6:** 315–320.

Tsukada Y., Fang J., Erdjument-Bromage H., Warren M.E., Borchers C.H., Tempst P., and Zhang Y. 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439:** 811–816.

Turner B.M. 2000. Histone acetylation and an epigenetic code. *Bioessays* **22:** 836–45.

Vakoc C.R., Mandat S.A., Olenchock B.A., and Blobel G.A. 2005. Histone H3 lysine 9 methylation and HP1γ are associated with transcription elongation through mammalian chromatin. *Mol. Cell* **19:** 381–391.

Vettese-Dadey M., Grant P.A., Hebbes T.R., Crane-Robinson C., Allis C.D., and Workman J.L. 1996. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *EMBO J.* **15:** 2508–2518.

Volpe T.A., Kidner C., Hall I.M., Teng G., Grewal S.I., and Martienssen R.A. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297:** 1833–1837.

Wang H., Wang L., Erdjument-Bromage H., Vidal M., Tempst P., Jones R.S., and Zhang Y. 2004. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431:** 873–878.

Wang H., Huang Z.Q., Xia L., Feng Q., Erdjument-Bromage H., Strahl B.D., Briggs S.D., Allis C.D., Wong J., Tempst P., and Zhang Y. 2001. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* **293:** 853–857.

Wang Y., Wysocka J., Sayegh J., Lee Y.H., Perlin J.R., Leonelli L., Sonbuchner L.S., McDonald C.H., Cook R.G., Dou Y., et al. 2004. Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* **306:** 279–283.

Whetstine J.R., Nottke A., Lan F., Huarte M., Smolikov S., Chen Z., Spooner E., Li E., Zhang G., Colaiacovo M., and Shi Y. 2006. Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell* **125:** 467–481.

Wood A., Krogan N.J., Dover J., Schneider J., Heidt J., Boateng M.A., Dean K., Golshani A., Zhang Y., Greenblatt J.F., et al. 2003. Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol. Cell* **11:** 267–274.

Workman J.L. and Roeder R.G. 1987. Binding of transcription factor TFIID to the major late promoter during in vitro nucleosome assembly potentiates subsequent initiation by RNA polymerase II. *Cell* **51:** 613–622.

Wysocka J., Allis C.D., and Coonrod S. 2006a. Histone arginine methylation and its dynamic regulation. *Front. Biosci.* **11:** 344–355.

Yamane K., Toumazou C., Tsukada Y., Erdjument-Bromage H., Tempst P., Wong J., and Zhang Y. 2006. JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell* **125:** 483–495.

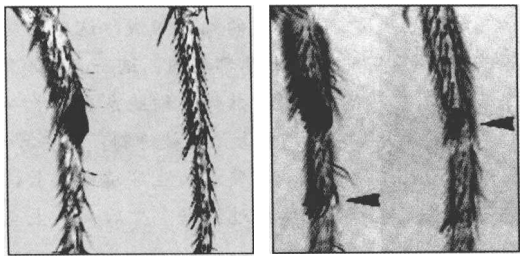
Yang X.J. and Seto E. 2003. Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. *Curr. Opin. Genet. Dev.* **13:** 143–153.

Zhang H., Roberts D.N., and Cairns B.R. 2005. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123:** 219–231.

Zhang Y. and Reinberg D. 2001. Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15:** 2343–2360.

Zhu B., Zheng Y., Pham A.D., Mandal S.S., Erdjument-Bromage H., Tempst P., and Reinberg D. 2005. Monoubiquitination of human histone H2B: The factors involved and their roles in HOX gene regulation. *Mol. Cell* **20:** 601–611.

# 第 11 章



# Polycomb 蛋白家族调控的转录沉默

Ueli Grossniklaus[1] and Renato Paro[2]

[1] *Institute of Plant Biology and Zürich-Basel Plant Science Center, University of Zürich, CH-8008 Zürich, Switzerland*

[2] *ZMBH, University of Heidelberg, D-69120 Heidelberg, Germany*

## 本章目录

1. 前言
   1.1 细胞记忆的概念
   1.2 Polycomb 家族的遗传学发现
2. 染色质沉默标记的建立
   2.1 PRC2 的组分及其进化保守性
   2.2 PRC2 的染色质修饰活性
   2.3 发育过程中 PRC2 的动态功能
3. 转录沉默的维持
   3.1 PRC1 的组分
   3.2 PRC1 如何靶向被沉默基因
   3.3 PRC1 抑制功能的建立
   3.4 “抗”沉默机制如何阻止遗传性抑制
4. 哺乳动物发育中的 PcG 抑制
   4.1 从基因到染色体抑制
   4.2 异常转录激活的后果
   4.3 维持干细胞命运
5. 总结及展望
参考文献

题图经许可，引自 Yamamoto et al. 1997

## 概要

人类、动物以及植物的器官都由大量不同类型的细胞组成，这些细胞行使着特殊的生理或结构功能。所有类型的细胞，其DNA编码的遗传信息都相同，这一点只有很少的例外。因此，只有通过基因特异的表达程序来产生特定类型的细胞。然而，细胞系在生长和细胞分裂过程中都必须保持一定的基因表达程序，也就是说，哪些基因被激活或被抑制，需要有一个记忆系统保证这个信息从母代到子代忠实传递。动植物人工培养组织即使在陌生环境中生长，也常常保持着分化的特征，这一点证明了这种系统的存在。例如常春藤经过组培重生后会产生与原始组织生长阶段相应的叶片类型（即新叶或成熟叶子）。

找出参与多次细胞分裂中维持一定细胞状态的分子机制（此过程被称为“细胞记忆”或“转录记忆”）是本章及后续几章所述的主要问题。同源异型框基因（*homeobox*，*HOX*）记忆能决定果蝇个体体节，对它的遗传分析发现了维持体节发育命运的重要调控因子。雄性果蝇的第一胸节的腿有性梳，第二、三胸节的腿则无此结构（见题图左）。20世纪40年代，人们发现有的雄性果蝇突变体所有腿上都有性梳（多聚梳和超量性梳）（见题图右），这是因为第二、三条腿转化成了第一条腿。遗传和分子的研究显示出这些突变影响的不是*HOX*基因产物本身，而是空间调控*HOX*记忆活性。多年以来，发现了很多这样的调控基因，它们主要分为两个相拮抗的类别，多聚梳类（Polycomb group，PcG）和三胸类（trithorax group，trxG），其中PcG蛋白为维持发育调控子（如*HOX*基因）的沉默所必需，而trxG蛋白一般对维持基因表达的激活状态有用。因此，PcG和trxG蛋白形成了细胞记忆的分子基础。

PcG和trxG蛋白组织形成大的多体复合物，通过调控染色质结构作用于其目的基因。本章我们关注于两个主要的多聚梳抑制性复合体（Polycomb repressive complex，PRC）PRC1和PRC2的分子实质，而trxG复合体的分子实质将在下章讨论。在果蝇中，PcG复合体通过PcG应答元件（PcG response element，PRE，一段DNA序列）被募集到靶基因上，然后建立起染色质的沉默，这种状态可以在多次细胞分裂中仍然保持。植物和动物的PRC2成员是保守的，而PRC1蛋白只存在于果蝇和脊椎动物，也就是说细胞记忆系统的基本元件既具保守性又有多样性；PcG复合体除了维持细胞类型还可能在干细胞可塑性中起到重要作用。在脊椎动物中，PcG一旦发生异常调节，则会导致肿瘤转移和癌症。因此对于多细胞真核生物来说，PcG蛋白对于许多正常发育和致病的基本过程都是至关重要的。

# 1. 前　言

所有多细胞生物体都起源于单个细胞——合子，合子继续发育，形成大量有着各自独特功能的细胞类型。这提出了一个问题：细胞类型被决定之后，如何在生长阶段的多次细胞分裂后依然保持呢？

## 1.1　细胞记忆的概念

一个成年动物有二、三百种不同结构的细胞类型，而植物有三、四十种，由于基因表达形式的复杂性，可区分出更多不同细胞类群。各种激活或沉默的特征基因表达形式，决定了每种细胞类型的身份和功能。在发育过程中和成年平衡状态时，需要在每次细胞分裂后忠实地记忆、重建这种基因表达状态，这一点很重要，尤其在有丝分裂中遗传物质复制（S 期）以及染色体分离过程（M 期）中，因为这二者在每个细胞周期都会打断基因表达。那么，细胞的不同基因表达形式是如何代代相传的呢？

20 世纪六、七十年代的实验表明，即使经过了长期的传代培养，动植物组织仍然会“记忆”一个确定的基因表达状态（Hadorn 1968；Hackett et al. 1987）。Hadorn 及其同事们指出，果蝇幼虫的成虫盘具有可遗传的细胞记忆，所以它们能记住在早期胚胎形成时所确定的表达状态。成虫盘是成簇的上皮细胞，它们保留在发育的胚胎中，是蜕变过程所产生的特异性外部结构和附属结构的前体。例如，第二胸节的两对成虫盘，其中一个形成一条中腿，另一个则形成翅膀（见图 12-2）。将成虫盘植入成年雌蝇的血腔中进行培育，它们就继续增殖而不分化。在蜕变之前，如果将其重新植回幼虫体内，它还会接着分化成预期的成体结构，即使之前经过了连续的移植传代也是如此。新的发现说明 PcG 和 trxG 蛋白也参与维持成虫盘细胞的先决状态（图 11-1），然而在以上的培育实验中，也有很少的成虫盘改变了发育命运，称为细胞的转决定（transdetermination）。此过程包含了转决定细胞中由 JNK 级联信号通路引起的 PcG 蛋白抑制的下游调控（图 11-2a，b）（Lee et al. 2005）。PcG 突变体转决定的频率增加，这也支持了 PcG 蛋白在维持成虫盘细胞命运上有作用的说法（Klebes et al. 2005）。这样看来，不管是正常发育中，还是再生过程中（详见 4.3 节），PcG 蛋白都应该在细胞命运重编程中起到重要作用。

## 1.2　Polycomb 家族的遗传学发现

在所有多细胞动物中，前后体轴都由 *HOX* 基因的特定表达形式决定（见图 12-2）。果蝇的胚胎发生过程中，母源（由卵子继承而来）转录因子与合子产生的转录因子形成了对应每个体节所需要的特异性 *HOX* 表达组合，而后尽管早期转录调控因子早已不存在，这个体节特异的 *HOX* 基因活性状态在果蝇整个发育过程中能够一直维持。*HOX* 基因功能在遗传学上被确定时，已经分离出很多反式调控因子，其中 *Polycomb*(*Pc*) 是由 Pam 和 Ed Lewis 发现的，他们还对其进行了遗传学分析（Lewis 1978）。杂合的雄性 *Pc* 突变体在第二、三条腿上有多余的性梳，而野生型雄性只在第一条腿上有（见题图）。纯合突变体是胚胎致死的，其表皮节均转化为最尾端的腹节（图 11-2 c，d）。这些经典的 PcG 表型被

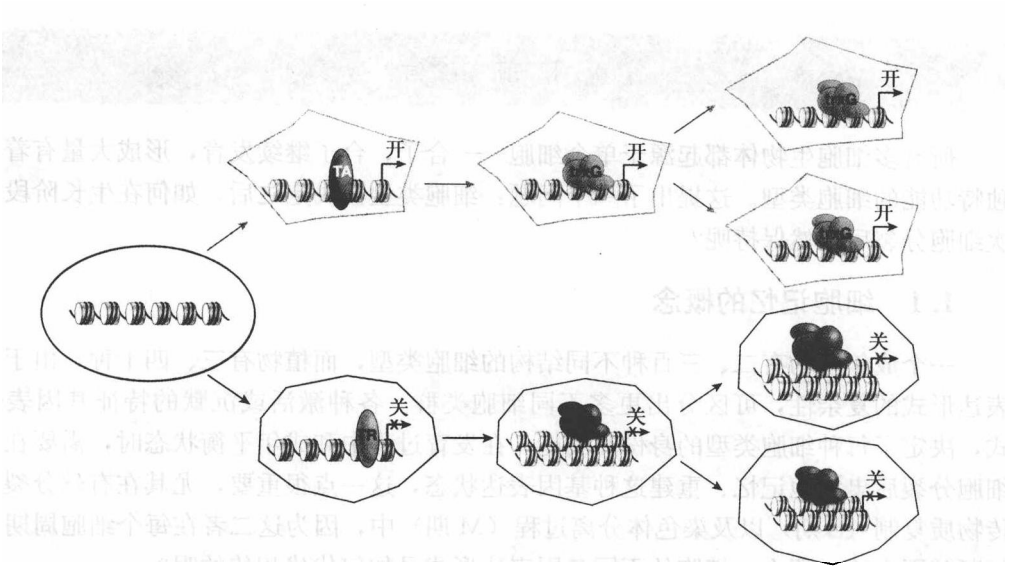



图 11-1 细胞记忆的概念

示意图：PcG 和 trxG 蛋白复合体决定基因表达处于激活或抑制状态，由此决定细胞分化，并在后续多次细胞分裂后保持不变。（TA）转录激活子；（TR）转录抑制子。

解释为由 *HOX* 基因的易位表达引起。因此，*Pc* 和其他有类似表型的基因都被定义为 *HOX* 基因的抑制子。后续更深入的分析发现，PcG 蛋白只能维持 *HOX* 的抑制，而不会影响 *HOX* 表达形式形成过程中建立位置特异性。后者是由早期作用于分节基因编码的转录因子完成的。这些新发现的反式调控子有的抑制 *HOX* 基因表达，有的激活表达，被分为两种拮抗类别，分别为 PcG 和 trxG 组（图 11-1）（Kennison 1995）。

鉴别分离出果蝇 PcG 基因分子之后，可以用小鼠作为模式生物研究它的脊椎动物同源物的功能，已经发现在小鼠中 PcG 也是 *HOX* 基因表达的关键调控子（van der Lugt et al. 1994；Core et al. 1997）。哺乳动物中，PcG 基因的突变会导致脊椎出现同源异型转化（图 11-2 e，f）。另外，PcG 基因在细胞增殖控制、干细胞维持以及癌症中起到重要作用（见第 4.2、4.3 节）。果蝇和哺乳动物的 PcG 基因之间惊人的保守性促进了生化分析，并因此促进了 PcG 复合体的另一些新成员的发现，如 RING1 蛋白（Satijn and Otte 1999）。例如，将小鼠中 RING1a 进行定向突变会导致经典的同源异型转化表型，之后才发现这也与果蝇中 PcG 基因超量性梳相对应。

通过对多个遗传学筛查找到的突变体的分子定性，发现了线虫和拟南芥的 PcG 同源蛋白。线虫有一种“母源不育效应”（*maternal effect sterile*，*mes*）突变体，在筛查此突变体时发现了 PcG 成员，它在雌雄同体的生殖细胞系中则参与沉默 X 染色体（Fong et al. 2002；见第 15 章）。

拟南芥的 PcG 基因是在研究不同发育进程的遗传学筛查中发现的（Hsieh et al. 2003）。植物中第一个 PcG 基因——卷叶基因（*CURLY LEAF*，*CLF*），其突变体发生了花器官的同源异型转化（图 11-2g、h）（Goodrich et al. 1997）。非受精依赖性种子基因（*Fertilization-Independent Seed*，*FIS*）的突变会导致母源效应的败育（Grossni-

klaus et al. 1998)，也可会使种子发育的一些情况在未受精时就可产生（Luo et al. 1999；Ohad et al. 1999)。最后，PcG 基因还在花期突变体中发现，例如，有的突变体在发芽后直接开花（Yoshida et al. 2001)，有的则被破坏了春化反应性（春化是使植物在长期对抗寒冷之后形成开花能力的过程；Gendall et al. 2001)。

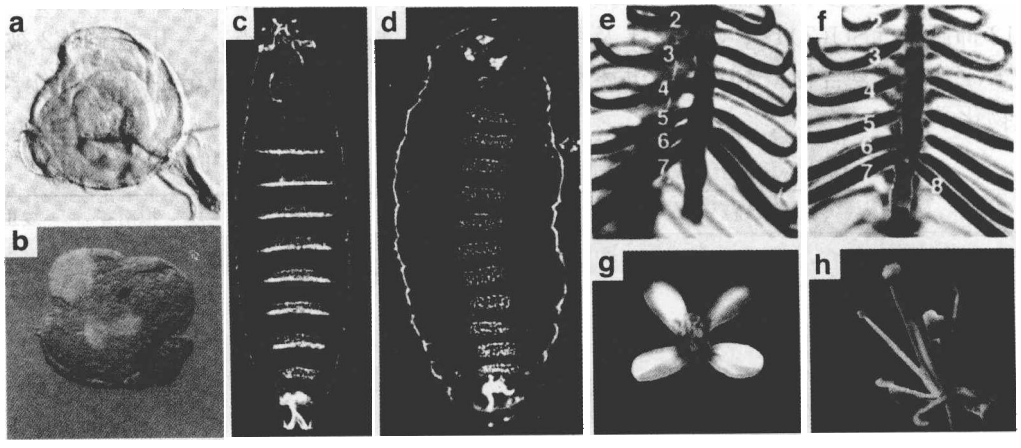



图 11-2　不同物种 PcG 突变体的同源异型转化

(a-d) 果蝇；(e，f) 小鼠；(g，h) 拟南芥；(a，b) 由翅膀特异性残翅基因 *vestigial* 的表达（GFP 标记)，可以看出腿成虫盘被转决定。(c，d) 野生型 (c) 和 *Su(z)12* 突变体 (d) 胚胎的表皮 (cuticles)。*Su(z)12* 突变体的所有腹节、胸节和几个头节（此聚焦层不可见）同源异型转化为多拷贝的第八腹节，这是因为 *Abd-B* 基因在这些体节中表达不正确。(e，f) 新生小鼠野生型 (e) 和 $Ring1A^{-/-}$ (f) 中轴骨架。清晰可见胸部骨架，骨（红色）及关节（蓝色)。突变体前部 (anterior) 转化为第八根胸椎，这里出现了八根 (1-8) 椎胸肋，野生型只有七根 (1-7)。(g，h) 野生型 (g) 以及 *clf-2* 突变型花朵。野生型花朵的萼片、花瓣、雄蕊和心皮排列正常，而 *clf-2* 花朵中，花瓣减少或消失 (a，b：由 Lee 和 Raro 提供；c，d 经许可转印自 Birve et al. 2001；e，f 经许可转印自 del Mar Lorente et al. 2000；由 Goodrich 提供)。

PcG 调控的多种生物学进程都说明了不同生物体中维持关键的发育调控子的抑制状态有多么重要。一方面，动植物之间的生物学功能有着惊人的保守性，例如，如同源异型基因，它们在动植物中都要调控关键的发育调控因子，而且这些 PcG 在细胞增殖的严格调控中都起到作用。另一方面，PcG 复合体是很灵活具有动态性的分子模块，可以调控多种发育和细胞的生物学过程。

## 2. 染色质沉默标记的建立

PcG 蛋白在生化上有两个分类：多聚梳抑制性复合体 1、2(PRC1 和 PRC2)，它们参与抑制基因表达的一连串步骤；首先，PRC2 具有组蛋白修饰活性，可将基因沉默的靶基因上 H3K27 和/或 H3K9 甲基化。然后 PRC1 复合体识别并结合到修饰处，引起染色质适当结构变化。PRC2 蛋白在所有多细胞模式生物中都存在，而 PRC1 复合体在线虫和拟南芥中都还没有发现。

## 2.1 PRC2 的组分及其进化保守性

人们已经从果蝇胚胎中纯化出几种 PRC2 的变异体，所有这些复合体都含有四种核心蛋白（Levine et al. 2004）：SET 组蛋白甲基转移酶 Zeste 基因增强子 E(Z)、含 WD40 结构域的多余性梳蛋白（extra sex comb，ESC）、组蛋白结合子 p55 以及 Zeste12 抑制子［SU(Z)12］（表 11-1 及图 11-3）。由于这种组成特点，一开始 PRC2 被称做 E(Z)-ESC 复合物。本节重点讲述不同模式生物中已知 PRC2 各成分的分子生物学和生物化学。

**表 11-1 模式系统中的核心 PcG 基因**

| | 果蝇 | | 小鼠 | 拟南芥 | 线虫 |
|---|---|---|---|---|---|
| PcG DNA 结合蛋白 | | | | | |
| *pho* | 多同源异型（Pleiohomeotic） | 锌指 | *Yy1* | | |
| *phol* | 类多同源异型(Pleiohomeotic-like) | 锌指 | | | |
| *Psq* | 小人物（Pipsqueak） | BTB-POZ 结构域 | | | |
| *Dspl* | 背部开关蛋白 1(Dorsal Switch protein1) | HMG 结构域蛋白 | *HMGB2* | | |
| *PRC2* 核心复合体 | | | | | |
| *esc* | 多余性梳（Extra sex combs） | WD40 重复片段 | *Eed* | *FIE* | *MES-6* |
| *E(z)* | zeste 增强子（Enhancer of zeste） | SET 结构域 | *Ezh1/Enx2* | *CLF* | |
| | | | *Ezh2/Enx1* | *MEA* | *MES-2* |
| | | | | *SWN* | |
| *Su(z)12* | zeste12 抑制子(Suppressor of zeste 12) | 锌指 | | *FIS2* | |
| | | VEFS 盒 | *mSU(Z)12* | | *VRN2* |
| | | | | *EMF2* | |
| *p55* | p55 | 组蛋白结合域 | *RbAp48* | *MSI1*（*MSI2*、 | |
| | | | *RbAp46* | *3*、*4*、*5*） | |
| PRC1 核心复合体 | | | | | |
| *Pc* | 多聚梳（Polycomb） | 染色质结构域 | *Cbx2/M33* | | |
| | | | *Cbx4/MPc2* | | |
| | | | *Cbx6* | | |
| | | | *Cbx8/MPc3* | | |
| | | | *Cbx7* | | |
| *Ph* | 多同源异型（Polyhomeotic） | 锌指 | Edr1/Mph1/Rae28 | | *SOP-2* |
| | | SAM/SPM 结构域 | *Edr2 /Mph2* | | |
| *Psc* | 尾部性梳（Posterior Sex Combs） | 锌指 | *Bmi1* | | |
| | | HTH 结构域 | *Rnf110/Zfp144/* | | |
| | | | *Mel18* | | |
| *dRing/Sce* | d 环/多余性梳(dRing/Sex combs extra) | RING 锌指 | *Ring1/Ring1a* | | |
| | | | *Rnf2/Ring1b* | | |

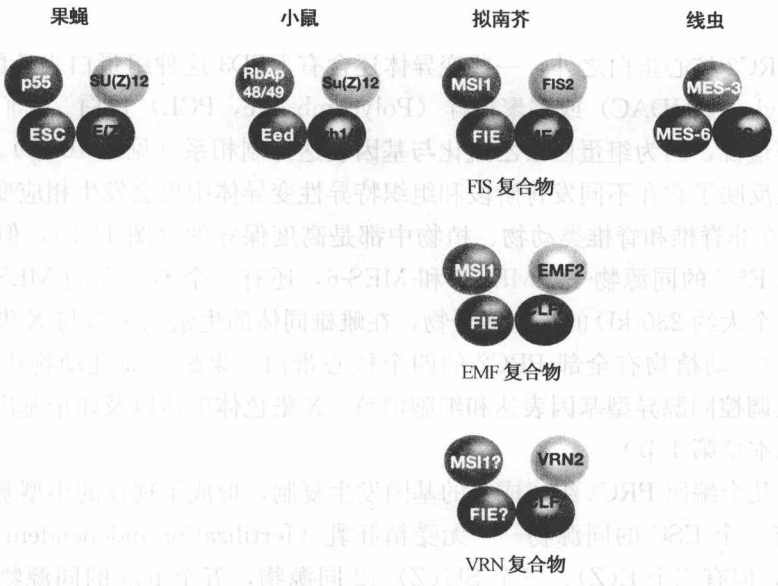



图 11-3　保守的 PRC2 核心复合物

果蝇、小鼠、拟南芥和线虫的 PRC2 核心成员。人们认为拟南芥的一个原始复合物在发育过程中分化成三个变异体，分别有着各自的功能。线虫的 PRC2 核心复合物只含有三种蛋白：MES-3 和其他已知的 PRC2 蛋白没有任何同源性。颜色代表同源性，接触代表相互作用（选自 Reyes and Grossniklaus 2003；Chanvivattana et al. 2004）。

*E*(*z*) 基因编码一个含有 760 个氨基酸残基的蛋白，它含有一个具有组蛋白赖氨酸甲基转移酶（histone lysine methyltransferase，HKMT）活性的 SET 结构域，前面连接着一个 CXC 或 PreSET 结构域（Tschiersch et al. 1994），这个结构域含有九个保守的半胱氨酸，它们结合三个锌离子，可以稳定 SET 结构域。它的这个结构性作用受以下事实的支持，几个温度敏感的 E(z) 等位基因影响这其中的一个半胱氨酸（Carrington and Jones 1996）；另外，E(z) 含有 SANT 结构域，这可能参与组蛋白结合以及 C5 结构域，它在与 SU(Z) 12 进行物理上的相互作用时是必需的。

ESC 是一种短小的蛋白，含 425 个氨基酸，其中有五个 WD40 重复片段，形成一个 β 螺旋桨结构，也是蛋白进行相互作用的平台，因此 ESC 在 PRC2 中起到中心的作用，它在所有分析的模式系统中都可以物理地与 E(z) 和 p55 作用。

SU(Z)12 蛋白含 900 个氨基酸，典型特征为 $C_2H_2$ 型锌指以及一个羧基端 VEFS 结构域，此结构域在 SU(Z)12 及其植物中三个同源物——VRN2、EMF2、FIS2 中都保守（图 11-3）。几个 *Su*(*z*)*12* 等位基因突变中此区域被改变，证明了它在与 E(Z) 的 C5 结构域作用中也是很重要的（Chanvivattana et al. 2004；Yamamoto et al. 2004）。

p55 蛋白不是通过遗传学手段分析 PcG 成员时发现的，很可能是因为它参与了很多其他染色质相关的蛋白复合物（Hennig et al. 2005）。它是 PRC2 蛋白的一部分，通过生化手段发现，这是因为 p55 含 430 个氨基酸，含有六个 WD40 重复片段，能够与

ESC 或其哺乳类、植物中同源物进行物理上的相互作用（Tie et al. 2001；Köhler et al. 2003a）。

除了 PRC2 核心蛋白之外，一些变异体还含有 RPD3 这种组蛋白去乙酰化酶（histone deacetylase，HDAC）或多聚梳样（Polycomb-like，PCL）蛋白。它们与 RPD3 的作用很值得重视，因为组蛋白去乙酰化与基因表达抑制相系（见第 10 章）。PRC2 的不同组成可能反映了它在不同发育阶段和组织特异性变异体中也会发生相应变化。

PRC2 在非脊椎和脊椎类动物、植物中都是高度保守的（图 11-3），但是线虫中只有 E(Z) 和 ESC 的同源物——MES-2 和 MES-6，还有一个不保守的 MES-3 蛋白，它们形成了一个大约 230 kD 的小型复合物，在雌雄同体的生殖系中参与 X 染色体的沉默（见第 15 章）。动植物有全部 PRC2 的四个核心蛋白。果蝇、哺乳动物中的复合体约 600 kD，在调控同源异型基因表达和细胞增殖、X 染色体失活以及印记基因表达中都有作用（详见本章第 4 节）。

植物中几个编码 PRC2 组成成分的基因发生复制，形成了现在的小型基因家族。拟南芥中只有一个 ESC 的同源物——无受精胚乳（fertilization-independent endosperm，FIE）蛋白，但有三个 E(Z)、三个 SU(Z) 12 同源物，五个 p55 的同源物（分别称作 MSI1～5）（表 1）。这些蛋白的不同组合至少形成了三个不同的复合体，分别控制特定的发育过程（图 11-3、11-4）（Reyes and Grossniklaus 2003；Chanvivattana et al. 2004）。

这些复合体中，无受精种子蛋白基因（*FIS*）家族研究得最为透彻，它们在控制种子细胞增殖过程中起到重要作用（Grossniklaus et al. 2001）。这个 FIS 或 MEA-FIE 复合体含有 MEDEA、FIE、FIS2 和 MSI1 蛋白。PHERES1 蛋白(PHE1) 为 MADS 结构域转录因子，FIS 复合体能调控编码 PHERES1 的基因；酵母细胞增殖调控中起重要作用的 Skp1，其同源物 MEIDOS 的基因的表达也受 FIS 复合体的调控（Köhler et al. 2003b）。有趣的是，*PHE1* 的父源等位基因表达量高于母源等位基因，这是由于 FIS 复合体特异地抑制了母源等位基因而控制了基因组印记（Köhler et al. 2005）。因此下面将要概括 FIS 复合体与哺乳类对应复合物在调控细胞增殖和印记基因表达上共有的功能。

EMF 复合体含有 CLF 及 EMBRYONIC FLOWER2(EMF2) 蛋白（Chanvivattana et al. 2004）。这二者之一如果发生了突变，则会显示出微弱的同源异型转化以及早花的表型。同源异型基因的联合作用决定了花器官的属性，EMF 复合体可以抑制这些基因（Goodrich et al. 1997）。因此，与果蝇和脊椎动物中的 *PRC2* 基因类似，EMF 有维持同源异型基因受抑制的功能（图 11-2）。然而，植物中的同源异型基因并不编码同源结构域蛋白，而是编码一些其他的转录因子，这些因子属于 MADS 结构域以及植物特异性 AP2 结构域家族。很强的 *EMF2* 突变体，其幼苗不经过发育的营养阶段就在萌芽后直接开花，这时却显示出非常严重的表型（Yoshida et al. 2001）。因此，EMF 复合体能阻止过早开花，在早期发育和后来花器官形成中都起到作用（Chanvivattana et al. 2004）。两个阶段的 EMF 复合体都抑制了花同源异型基因如 *AG*

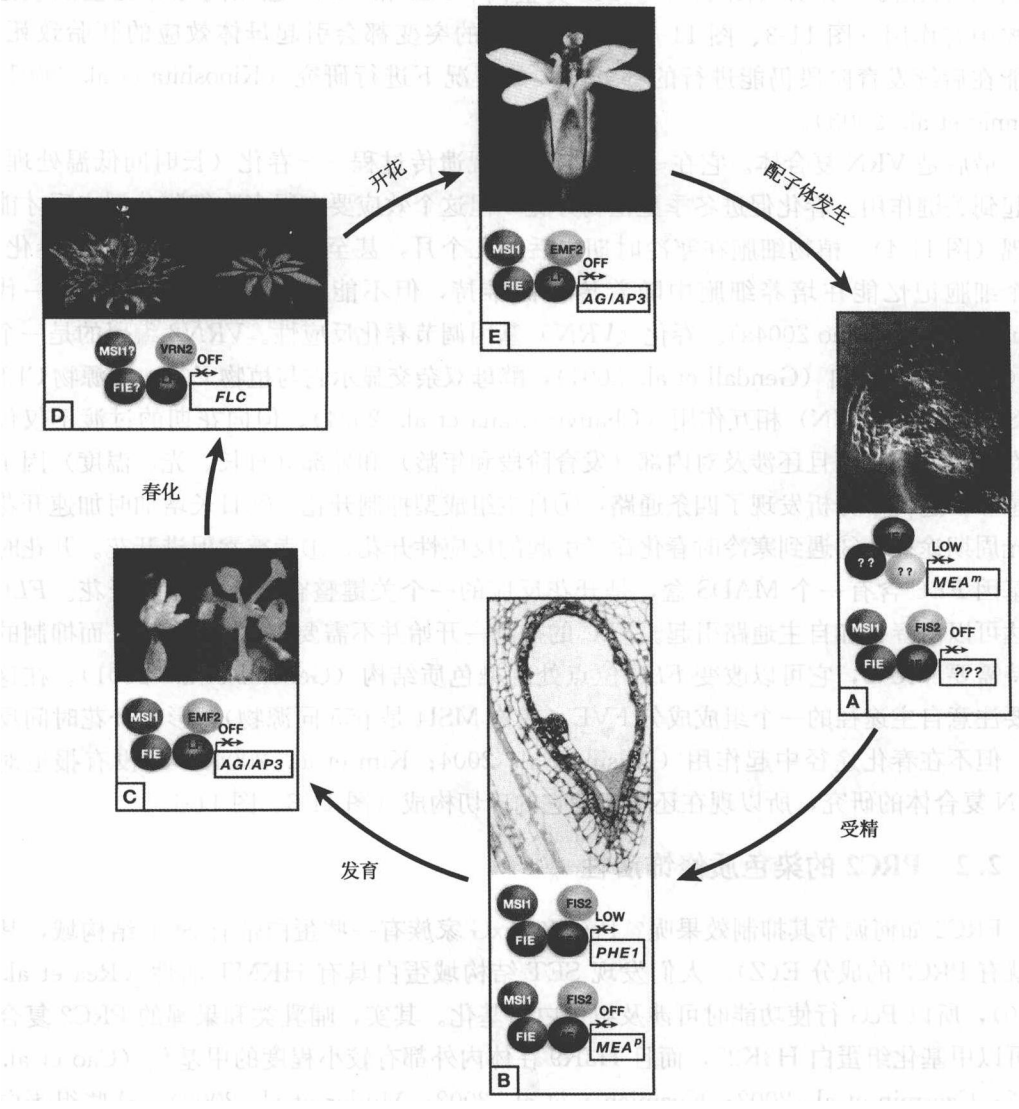



图 11-4　不同 PRC2 在植物发育不同阶段的作用

植物的生活史中，PRC2 的不同变异体（图 11-3）控制了发育过程。（A）野生型胚珠中心为雌配子体清晰可见。FIS 复合体抑制了控制中心细胞扩增的基因；在所有 *fis* 类突变体中，细胞增殖可以在不受精的情况下进行。受精时，MEA 可以维持 $MEA^m$（母源）的低表达，而这个活性不需要依赖其他 FIS 复合体的成分。（B）野生种子含有胚和胚乳，外面包有种皮。受精后，FIS 复合体在胚和胚乳细胞扩增的调控中起作用，它能维持 *PHE1* 的低表达，用来保持 $MEA^p$（父源）等位基因的沉默。（C）萌芽 21 天后的野生型（右）和 *emf2* 突变体（左）幼苗。*emf2* 幼苗产生了同源异型转化的花，但没有叶子。EMF 复合体阻止开花，抑制花同源异型基因，如 *AG*，*AP3* 等。（D）春化（右）和未春化（左）的植物，后者营养期长，产生大量叶子。发育的营养期，内外源信号诱导开花。春化过程使开花抑制子 *FLC* 作用减弱，从而促进开花。VRN 复合体能维持这个作用减弱的状态。（E）野生型拟南芥花。花的器官形成中，EMF 复合体调控了花同源异型基因，这些基因决定花器官的种类（A，J. M. Moore and U. Grossniklaus 提供；B，J. -P. Vielle-Calzada and U. Grossniklaus 提供；C，经许可转印自 Moon et al. 2003[© ASPB]；D，经许可转印自 Sung and Amasino 2004a；E，可转印自 Page and Grossniklaus 2002 [© Macmilan])。

和 *APETALA3*（AP3）（图 11-4）。FIS 类蛋白、FIE 和 MSI1 也在同源异型基因表达调控中有作用（图 11-3、图 11-4）。由于二者的突变都会引起母体效应的胚胎致死，只能在后续发育阶段仍能进行的部分缺失的情况下进行研究（Kinoshita et al. 2001；Hennig et al. 2003）。

最后是 VRN 复合体。它在一个著名的表观遗传过程——春化（长时间低温处理）中起到关键作用。春化促进冬季型植物开花，但这个效应要在很多次细胞分裂之后才能表现（图 11-4）。植物细胞在寒冷时期之后的几个月，甚至几年后还记得它已经春化。这个细胞记忆能在培养细胞中随着传代而维持，但不能从植物的一代传到下一代（Sung and Amasino 2004a）。春化（*VRN*）基因调节春化反应性。*VRN2* 编码的是一个 SU(Z)12 同源蛋白（Gendall et al. 2001），酵母双杂交显示它与植物 E(Z) 同源物 CLF 和 SWINGER(SWN) 相互作用（Chanvivattana et al. 2004）。但向花期的过渡不仅仅由春化来调控，而且还涉及对内部（发育阶段和年龄）和外部（日长、光、温度）因子的感知。遗传学分析发现了四条通路：①自主组成型抑制开花；②日长增加时加速开花的光周期途径；③遇到寒冷时春化途径引起的反应性开花；④赤霉素促进开花。开花时间基因 *FLC* 含有一个 MADS 盒，是开花反应的一个关键整合子，能抑制开花。*FLC* 表达可以由春化和自主通路引起。*FLC* 的抑制一开始并不需要 VRN 复合体，而抑制的维持需要 VRN2，它可以改变 *FLC* 位点处的染色质结构（Gendall et al. 2001）。在这里要注意自主途径的一个组成成分 FVE（或称 MSI4 是 p55 同源物），影响开花时间反应，但不在春化途径中起作用（Ausin et al. 2004；Kim et al. 2004），还没有报道对 VRN 复合体的研究，所以现在还不清楚它的确切构成（图 11-3、图 11-4）。

## 2.2 PRC2 的染色质修饰活性

PRC2 如何调节其抑制效果呢？PcG 和 trxG 家族有一些蛋白带有 SET 结构域，其中就有 PRC2 的成分 E(Z)。人们发现 SET 结构域蛋白具有 HKMT 活性（Rea et al. 2000），所以 PcG 行使功能时可涉及组蛋白甲基化。其实，哺乳类和果蝇的 PRC2 复合体可以甲基化组蛋白 H3K27，而且 H3K9 在体内外都有较小程度的甲基化（Cao et al. 2002；Czermin et al. 2002；Kuzmichev et al. 2002；Müller et al. 2002）。这些组蛋白标记一般与转录沉默状态相联系。进一步说，H3K9 和 H3K27 甲基化与 bithorax 复合体的同源异型基因控制有关（Müller et al. 2002）。然而，E(Z) 突变体中只缺少了 H3K27，可见其在 PcG 沉默中的重要性。SU(VAR)3-9 蛋白自身就能甲基化 H3K9，而 E(Z) 蛋白自己没有甲基化 H3K27 的 HKMT 活性，最小的能起 HKMT 作用的复合体也需要 ESC 和 SU(Z)12，这二者可能有调节功能。最近又发现 PRC2 复合体还可以甲基化 H1K26(Kuzmichev et al. 2004)。Eed 是哺乳类中 ESC 同源物，其不同异构体决定了哺乳类 PRC2 会进行 H1K26 还是 H3K27 甲基化（Kuzmichev et al. 2004）。然而 H1K26 甲基化与 PcG 沉默是否有功能上的联系目前还不知道。

植物中 PRC2 还没有证明在体外有 HKMT 活性，但 *FLC* 调控研究发现春化作用致使乙酰化丧失而 H3K9、H3K27 甲基化增加，这个增加的甲基化大多数位于此基因

的第一个内含子（Bastow et al. 2004；Sung and Amasino 2004b）。这两个甲基化标记在 *vrn2* 突变体中都消失了，说明是 VRN 复合体将这些组蛋白加上了抑制性的甲基化标记。在两个其他突变体 *vrn1* 和春化不敏感 3（*vernalization insensitive3*，*vin3*）中，只有 H3K9me2 标记消失。*VRN1* 和 *VIN3* 分别编码了 B3 结构域和同源结构域家族的转录因子，但它们参与染色质修饰的准确分子机制目前还不清楚。

迄今有大量研究表明 PRC2 的主要功能是与 HKMT 活性相关的，但是一些 PRC2 变异体还有其他的染色体修饰活性。*Rpd3* 基因编码一种 PcG 沉默中起作用的 HDAC（Tie et al. 2001）．可是尽管 *rpd3* 突变会加强 PcG 表型，它们却没有显示出典型的同源异型转化。并不是所有 PRC2 制备物中都含有 RPD3，也许是因为它与 PRC2 核心蛋白的作用整体上比较微弱，或可能由于这种作用有时空特异性。RPD3 和 PRC2 的相互作用是一个有意思的合作，因为 HKMT 和 HDAC 活性都与染色质沉默有关，二者相结合还会增强这个结果。

## 2.3　发育过程中 PRC2 的动态功能

如前所述，PRC1 和 PRC2 核心成分与不同因子相关，这些因子可能将 PcG 复合体募集到组织特异的靶位点或者调控其活性（Otte and Kwaks 2003）。PcG 抑制的不同阶段在图 11-5 中有说明，图 11-5 阐明了 PcG 复合体在果蝇胚胎形成中不同阶段的特异性组成。现在要确定果蝇不同组织或细胞类型之间这些复合体究竟有什么区别还很困难，因为常常是整个胚胎提取物被用作后续的生化纯化。研究哺乳类和植物却明确显示 PcG 复合体在特异组织中有不同组成，而且在细胞分化中其组分也发生变化（Chanvivattana et al. 2004；Kuzmichev et al. 2005；Baroux et al. 2006）。哺乳动物的 PcG 基因表达量在不同细胞系之间有很大差异，甚至相同细胞中调控不同靶基因的 PcG 复合体都有不同组成。所以它在不同发育阶段的作用是动态的。

果蝇的同源异型基因在早期胚胎形成时处于抑制状态，而 PcG 蛋白用来维持这个状态，由此决定发育。一旦 PcG 靶基因的沉默状态已经固定，就会在这个个体剩余全部生命过程中保持。植物的 VRN 复合体有类似的情况：一旦被春化，靶基因就永久失活了，只有到再下一代才能回到初始状态。其他植物 PRC2 复合体对发育或环境刺激反应的更快一点，例如，FIS 复合体的一个功能就是抑制无受精细胞增殖，而一旦受精则很快诱导细胞分裂，这是因为 PcG 的靶基因被去抑制了。这说明了 PcG 抑制的是默认的状态，要能够正常发育则需要一些未知机制来克服。PcG 复合物失活是正常植物生活史的一部分，这可能是植物中没有 PRC1 同源蛋白的原因（图 11-4），PRC1 在靶基因永久稳定且长效失活中起到重要作用，这会对植物发育有害，植物中的 PcG 抑制是在适当刺激下就能解除的。

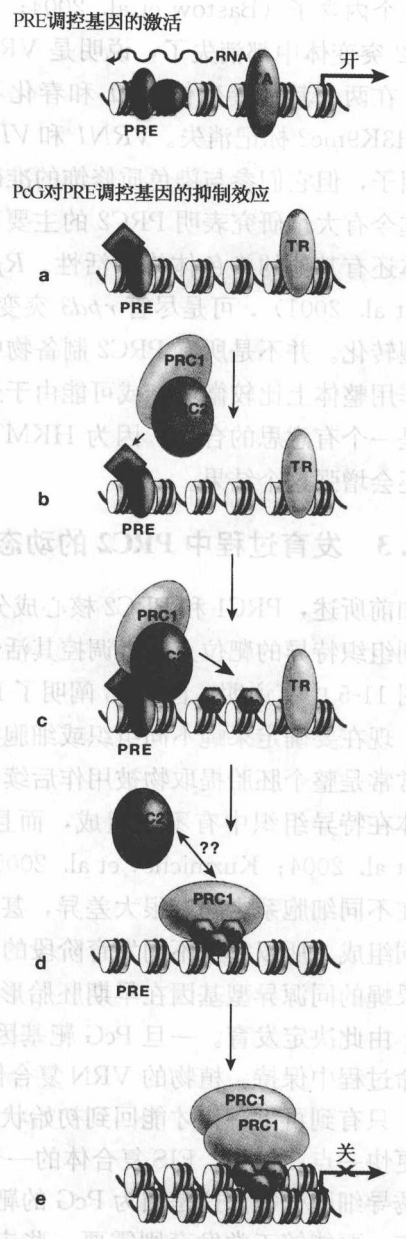



图 11-5 果蝇胚胎逐步产生 PcG 依赖的基因表达抑制的过程

PRE 调控基因的原始表达状态由转录调控子——转录抑制子（transcriptional repressor，TR）或激活子（transcripnal activator，TA）决定。转录通过 PRE 阻止“关闭”状态形成，而是促使了 trxG 依赖的“打开”状态（详见图 12-8）。（a、b）未转录的 PRE 结合特异性 DNA 结合蛋白（如 PHO、PHOL、DSP1 或 GAF），这些蛋白参与募集早期 PcG 复合体（含有 PRC1 和 PRC2）。（c）这个早期 PcG 复合体通过 E(Z) 依赖的组蛋白甲基化方式标记染色质。（d）当初始转录抑制子不在时，沉默状态就通过两个不同复合体——PRC1、PRC2 的相互作用来维持；PRC1 的 PC 染色体调节域与 H3K27me 3 结合使之稳定结合在上面。（e）PRC1 能使染色质压缩，进一步形成浓缩的沉默染色质。

## 3. 转录沉默的维持

### 3.1 PRC1 的组分

对 PcG 基因产物的分子分析发现了结构不同的染色质相关蛋白。PRC1 蛋白有四个 PcG 蛋白：Polycomb(PC)、多同源异型（PH）、尾部性梳（PSC）和 Ring1(dRing1/SCE)（表 11-1）(Francis et al. 2001)。这些蛋白以化学计量计，其他伙伴蛋白在不同

的纯化材料中发现。哺乳动物细胞中纯化出了一个相关复合体，意味着这四个亚基形成了 PRC1 的核心（Levine et al. 2002）。使用抗 PRC1 蛋白抗体对果蝇多线染色体的免疫荧光结果表明这些蛋白的定位有重叠，因此它们应该联合作用于一套确定的共有靶基因（图 11-6a）。另外，染色体上大约有 100 条带，所以 *HOX* 基因只是一个更大的调控网络中的一块，还有其他受 PcG 沉默作用的基因。

*PC* 基因编码了一个含 390 个氨基酸的蛋白质，N 端含有一个 Chromo 结构域，这个保守结构域与 HP1 有同源性，HP1 是果蝇异染色质形成所需要的蛋白质（Paro and Hogness 1991；见第 5 章）。之后又发现 Chromo 结构域还与 H3K27 及 H3K9 甲基化的部分相结合（Bannister et al. 2001；Fischle et al. 2003）。其 C 端有另一个保守结构域，而且如果该位置突变会导致一些异常，这些都说明蛋白的这个部分有着目前仍未知的调节功能。在靶向靶基因的过程中 PC 的 C 端是不必需的（由 Chromo 结构域完成），但是在体外实验中发现它与核小体作用（Breiling et al. 1999）。是否还有未发现的结构域能进行另一些组蛋白修饰还有待研究。人类 Pc2 只发现了一种 SUMO 的 E3 连接酶活性，所以 SUMO 修饰也是 PcG 沉默过程中的一个重要标记（Kagey et al. 2003）。

PSC 蛋白的 N 端部分在脊椎动物原始——原癌基因 *bmi-1* 和肿瘤抑制子基因 *mel-18* 之间是保守的，此区域含有一个 $C_3HC_4$ 环指结构域，可能调节了蛋白相互作用。*Bmi-1/Mel-18* 的亚核定位与细胞转化有关，而环指结构域在其中起到作用。

果蝇 *polyhomeotic*（*ph*）位点复制后就有近端（*ph-p*）和末端（*ph-d*）两个基因，它们有极大的同源性。小鼠的 PH 同源蛋白也有发现，所有这些蛋白都有一个保守的单个锌指以及一个 SAM（也称 SEP 或 SPM）结构域。这个结构域在其他 PcG 蛋白——第二腿性梳蛋白（sex combs on the midleg，SCM）上也发现了。在 PH 与其他同型/异型蛋白作用中有 SAM 结构域参与，可见它与蛋白相互作用有关。这些发现都支持了这样的观点：该蛋白能产生大的蛋白复合体参与沉默。其实 PcG 蛋白的亚核定位于 PcG 体，本身就可能是沉默的体系（Saurin et al. 1998）。

以上提到 dRING1 蛋白一开始不是作为 PcG 成员发现的，后来进行的生化纯化实验才发现这个因子在 PRC1 中和一个 RING 指结构域一同出现，它在其中可能起到结构性作用（Francis et al. 2001；Lavigne et al. 2004）。Ring1A、Ring1B 蛋白在哺乳动物 PRC1 中与失活 X 染色体上 H2A 泛素化有关，而且这个组蛋白标记的维持也需要 Ring1 蛋白（de Napoles et al. 2004；Fang et al. 2004；Cao et al. 2005；详见 4.1 节以及第 17 章）。

这四个蛋白组成了 PRC1 的核心结构，但复合体还与 SCM、Zeste 之类的 PcG 蛋白相联系（Otte and Kwaks 2003）。它们在 PRC1 中分子功能还不清楚，因为它们似乎是在核仁中发挥作用的，例如 Zeste 就有转录激活因子的功能。另外通过找 PcG 核心基因的转录调控子又发现了其他的 PcG 基因（Ali and Bender 2004）。三个 PcG 基因是在编码 PRC1 组分基因的上游调控子，PRC1 组分之间有一个负反馈的循环，而且受 PRC2 的正调控，进一步证明 PcG 基因之间有着复杂的交叉调控网络，以保证复合体中精确的各蛋白量（图 11-7a）。类似的，拟南芥中 FIS 复合体基因之间也有复杂的调控作用（Baroux et al. 2006）。

## 3.2 PRC1 如何靶向被沉默基因

对果蝇的同源异型基因簇进行分析得到了维持 *HOX* 基因适当体节特异性表达的调控元件，它们是 DNA 元件，叫做 *Polycomb* 应答元件（*Polycomb* responsive element，PRE），能在胚胎起始阶段之后维持这种表达。当 PRE 被整合到多线染色体的其他位点时可以募集 PRC1 中蛋白，说明它们可以决定 DNA 序列是否可以被 PRC1 蛋白识别和结合。然而 PcG 靶向的机制却很复杂。功能上确定的 PRE 可以从几百到几千 bp，有多种不同 DNA 结合蛋白的保守结合位点，而且一般在一个给定目标位点有两个 PRE 或者更多。迄今，所有确定的 PRE 都来自果蝇，没有在哺乳动物或植物中发现过。PRE 虽然有其复杂性，但其保守性序列区域确定有四个，都在果蝇 PRE 功能中起作用。其中一个（GCCAT）能与 PHO 和 PHOL 蛋白结合，这二者部分上有广泛的功能。PHO 和 PHOL 在 PcG 靶向时发挥功能，因为在早期胚胎提取物的 PcG 复合体中发现了二者，它们与 PRC1/2 的成员可以免疫共沉淀而且体外能结合 PRE(Poux et al. 2001)。近来还发现 DSP1 参与募集 PcG，它与许多 PRE 都有的 GAAA 结构域结合（Dejardin et al. 2005）。最后，trxG 蛋白 Zeste 以及 GAF（由 *Trithorax-like* 基因编码）可能有助于将 PcG 蛋白募集到目标基因上。

由于发现了有成对 GAF、Zeste 和 PHO/PHOL 的聚集位点是 PRE 的标志，所以新发展起来一种思路通过这一点可相对准确地预测已知 PRE 的存在，从而在果蝇基因组中找到新的 PcG 潜在靶基因（图 11-6 b）（Ringrose et al. 2003）。PRE 调控基因家族包括从赫赫有名的发育关键转录因子基因，到编码细胞周期调控和凋亡所涉及因子的基因，前者也是建立一定表达状态所必需的。

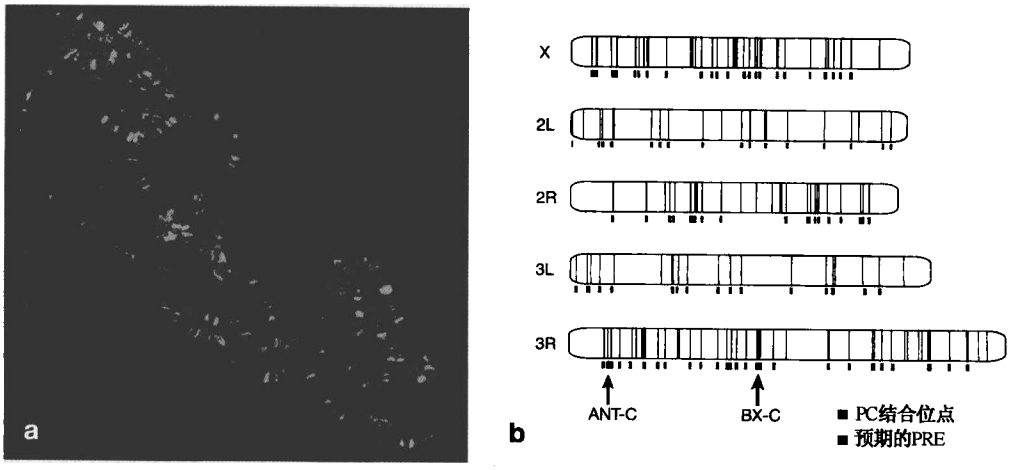



图 11-6 多线染色体上 PRC1 被定位至 PRE

(a) 对果蝇唾腺染色体进行免疫荧光染色可见 PC 蛋白分布 (b) 比对染色体各臂发现果蝇基因组的预期 PRE 位点与唾腺染色体细胞图谱上 PC 结合位点有重叠部分。两个 *HOX* 基因簇——*ANT-C*、*BX-C* 是 PRC1 的主要结合位点。

PRC1 一旦结合上去就和相邻组蛋白作用，在 PRE 处形成稳定的沉默复合体（图 11-7 b）。PRC2 产生 H3K27me3 标记是 PC 染色质调节域的另一个结合点（图 11-7 c）。没有这

些标记时，如可溶性甲基化组蛋白尾的肽段与之竞争，PRC1 就从靶基因上脱落下来（Czermin et al. 2002；Ringrose et al. 2004；Wang et al. 2004）。PRC2 有HKMT活性，而且产生发现沉默染色质典型的相关组蛋白标记，这可能说明了 PcG 抑制有一种新的机制。在 PRC2 催化 H3K27me3 形成之后，PRC1 通过 PC 蛋白的 Chromo 结构域结合到上面并稳定了沉默状态。以下发现也证明了这一点：①H3K27me3 标记和 PC 在多线染色体上有相同的定位；②E(z) 突变体中没有 PC 的结合。因为它缺少 HKMT 活性，也就不能将 H3 甲基化为 H3K27me3 标记，所以无法将 PC 募集到目标上（图 11-5）。尽管这个模型很吸引人，但 PRE 处的情况似乎更加复杂，因为 PRC2 和 PRC1 不是顺序作用的，而是在早期胚胎发生时共同作用于 PRE 上的（图 11-5 b，c）。因此，有可能 H3K27 甲

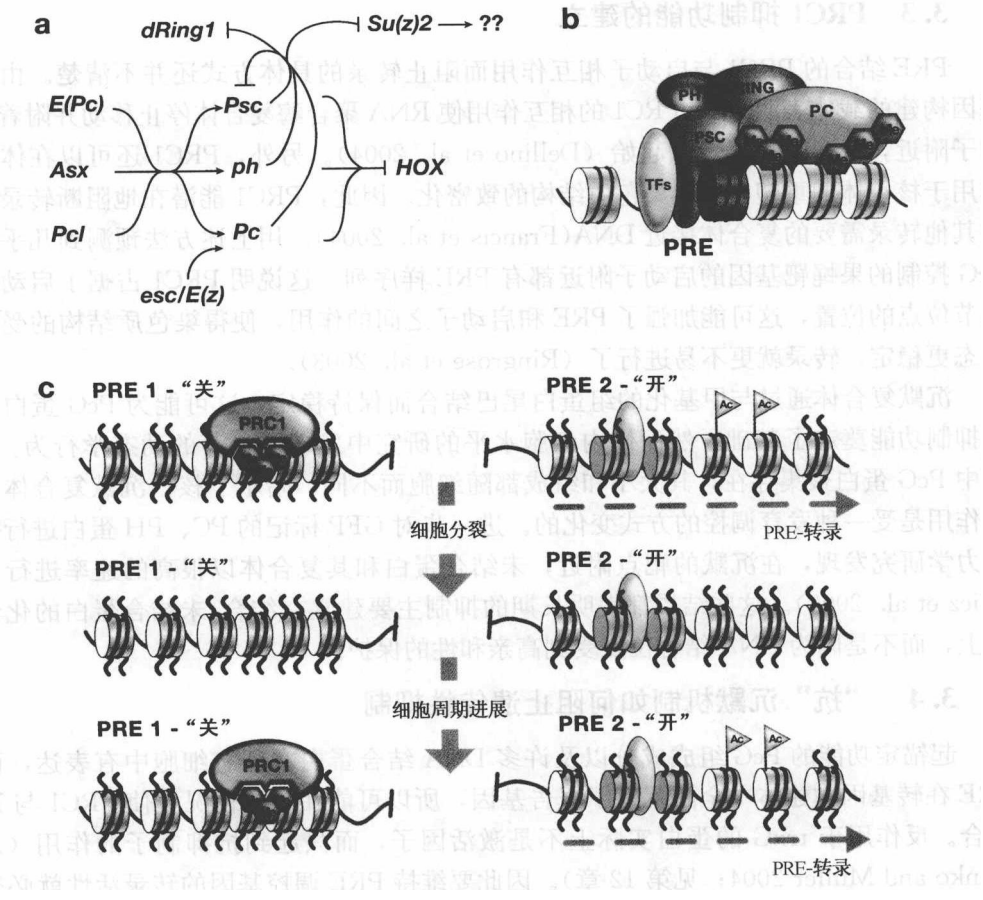



图 11-7　细胞分裂时 PRC1 调控及功能

(a) PcG 基因之间的交叉调控作用，遗传学研究为此提供了证明。E(Pc)、Pcl 和 Asx 都是核心 PRC1 成员上游作用的正调控子。PRC2 成员 *Esc* 和 *E*(z) 对 Pc 转录起到正调控作用。核心 PRC1 成员对 Psc、dRing1 以及 Su(z)12 有负反馈。基因产物得到精细的调控才能保证化学平衡，各种生物学过程才能很好的平衡。(b) 序列特异的转录因子将 PRC1 的各成分聚集到 PRE 上。PRC1 通过 PC 的染色质结构域锚定到相邻甲基化组蛋白尾巴上之后才能使沉默复合体得到稳定。(c) 差异化基因表达形式遗传的可能模型。基因间转录进程在控制活性基因的 PRE 处（PRE2）加上正的表观遗传标记（如乙酰化组蛋白尾巴、组蛋白变异体）。其他 PRE(PRE1) 在默认状态时被沉默。在 DNA 复制及有丝分裂过程中，只有正的表观遗传信号需要被传递到子细胞，保证下个细胞分裂间期基因间，在其他 PRE 位点默认被沉默之前，转录重新在 PRE2 处起始（a 摘自 Ali and Bender 2004）。

基化是 PcG 募集之后的下游事件，但在建立沉默状态时起到至关重要的作用。

以上描述的模型与异染色质形成相类似，异染色质蛋白 1(heterochromatin protein，HP1）就是通过其 Chromo 结构域与 SU(VAR)3-9 产生的 H3K9me 相作用而募集上来的（见第 5 章）。因此，转录因子将一个沉默复合体靶向到特定的 DNA 序列元件上，加上附近的带有合适修饰的组蛋白层，就产生更高级别的抑制染色质结构（图 11-5）。

在进化过程中，PRE 只有很少的序列保守性，即使很近的果蝇品系之间，PRE 的数量、位置和组成都有本质的差别（L. Ringrose and R. Paro，未发表）。这说明 PRE 的序列要求和位置是灵活多变的，可能与品系特异的要求相吻合。而 PRC1 的组分有高度保守性，被认为使用了相同的基础分子机制来诱导已沉默的靶基因产生更高级别染色质变化。

### 3.3 PRC1 抑制功能的建立

PRE 结合的 PRC1 与启动子相互作用而阻止转录的具体方式还并不清楚。由报告基因构建的载体表明 PRE-PRC1 的相互作用使 RNA 聚合酶复合体停止移动并附着在启动子附近，从而阻止了转录起始（Dellino et al. 2004）。另外，PRC1 还可以在体外反作用于核小体的重塑，诱导染色质结构的致密化。因此，PRC1 能潜在地阻断转录因子和其他转录需要的复合体接近 DNA(Francis et al. 2004)。用上述方法预测到几乎所有 PcG 控制的果蝇靶基因的启动子附近都有 PRE 样序列。这说明 PRC1 占据了启动子和调节位点的位置，这可能加强了 PRE 和启动子之间的作用，使得染色质结构的受抑制状态更稳定，转录就更不易进行了（Ringrose et al. 2003）。

沉默复合体通过与甲基化的组蛋白尾巴结合而保持稳定，这可能为 PcG 蛋白的长期抑制功能奠定了基础。然而体内细胞水平的研究中发现了显著的动态学行为。PcG 体中 PcG 蛋白聚集存在，其大小和组成都随细胞而不同，暗示了核仁沉默复合体的相互作用是受一种发育调控的方式变化的。进一步对 GFP 标记的 PC、PH 蛋白进行体内动力学研究发现，在沉默的靶点附近，未结合蛋白和其复合体以很高的速率进行交换(Ficz et al. 2005)。这些结果都表明长期的抑制主要建立在结合/未结合蛋白的化学平衡上，而不是因为 DNA 结合位点受到高亲和性的保护。

### 3.4 “抗”沉默机制如何阻止遗传性抑制

起锚定功能的 PcG 组成成分以及许多 DNA 结合蛋白在所有细胞中有表达，而且 PRE 在转基因构建体中全面沉默了报告基因，所以可能在默认情况下使 PRC1 与 PRE 结合。反作用于 trxG 的蛋白实际上不是激活因子，而只起到抗抑制子的作用（Klymenko and Müller 2004；见第 12 章）。因此要维持 PRE 调控基因的转录活性就必须以一种组织、发育阶段特异性的方式组织此 PRE 的沉默。例如在果蝇中，*HOX* 基因的激活就是由分节基因编码的转录因子早期级联而控制。有趣的是这些因子不仅诱导 *HOX* 基因，还诱导了基因间非编码 RNA 的转录，这是通过其上下游常有的相关 PRE 而完成的（图 11-5）。使用转基因构建体的研究表明，要阻止沉默，保持报告基因的活性状态必须有 PRE 参与的转录（Schmitt et al. 2005）。转录过程最可能的是将 PRE 染色体重塑，产生某一种确定的活性状态，如缺少抑制性的组蛋白甲基化以及产生组蛋白乙酰化。因此，即使 DNA 结合蛋白将 PRC1 募集到这一特定的已激活 PRE 上，组蛋白

的环境仍然不允许 PC 通过 H3K27me3 结合上来，因此也不能建立起稳定的沉默状态。既然 PcG 系统中是默认诱导沉默的，那么差异化基因表达的表观遗传只需要在 DNA 复制以及有丝分裂过程中活性的 PRE 得到传递（图 11-7c）。而这在分子水平上是如何达到的，哪些表观遗传标记负责维持活性 PRE 状态，都还是待解决的问题。有趣的是，最近发现在一个果蝇同源异型基因 *Ubx* 的 PRE 处，产生的非编码 RNA 一直结合于染色质并募集 trxG 调控子 ASH1(Absent Small or Homeotic discs1)，一旦这些 RNA 通过 RNAi 被破坏，募集到 PRE 上的 ASH1 也会减少，表明这个相互作用在同源异型基因的表观遗传激活上起到重要作用，它们通过克服了默认的 PcG 诱导的沉默（Sanchez-Elsner et al. 2006）。

## 4. 哺乳动物发育中的 PcG 抑制

### 4.1　从基因到染色体抑制

小鼠 PRC1 成员发生突变会显示出椎骨的同源异型转化，然后出现一块多余的脊椎，这是 *HOX* 基因去抑制的结果（图 11-2 e，f）(Core et al. 1997)。而且，突变小鼠的造血干细胞没有增生反应，造成了严重的联合免疫缺陷（Raaphorst 2005）。PcG 蛋白的作用在血细胞中得到了尤其充分地研究，因为大多数血细胞系的细胞类型特异的转录程序了解得很清楚。然而，细胞系的保证和限制性要以某种方式在细胞分裂时也忠实地维持。PcG 敲除小鼠的 B-T 细胞前体群都能正常产生，说明 PcG 控制不涉及建立细胞系特异性基因表达，PcG 蛋白不涉及决定某一特定发育通路之后的决定，而是参与到使细胞系选择不可逆转。

*HOX* 基因表达方式决定了不同血细胞系，除了控制这个基因，PcG 蛋白还在控制增殖中起到重要作用。*Bmi1* 基因是果蝇 *Psc* 的同源物，一开始被发现是原癌基因，在与 *myc* 联合作用时可以诱导小鼠的淋巴发生（van Lohuizen et al. 1991）。Bmi 蛋白控制细胞调节子 $p16^{INK4a}$ 和 $p19^{ARF}$ (Jacobs et al. 1999)。Bmi1 及其相关蛋白 Mel-18 都是 INK4c-ARF 位点的负调控因子，正常淋巴增殖控制需要这一位点，如果这一细胞周期重要检验点调控错误，会影响小鼠的凋亡和衰老（Akasaka et al. 2001）。

哺乳动物 PcG 蛋白还与经典的表观遗传现象——X 染色体失活（见第 17 章）有关。雌性 XX 细胞中一条 X 染色体失活，伴随了一系列染色质修饰，其中就有 PcG 蛋白参与（Heard 2004）。尤其是 Eed 或 E(Z) 同源物 Enx1（表 11-1），在转录沉默相关的组蛋白标记建立中起重要作用。这个 PRC2 临时结合到包被有 *Xist* RNA 的 X 染色体上，同时进行 H3K27 甲基化。相反，*eed* 突变小鼠胚胎没有 Enx1 HKMT 的募集，也没有发现任何 H3K27 甲基化。然而如果没有 PRC2 组分也不会导致整个失活 X 染色体的去抑制，而是 X 染色体上基因会出现散发性的重新表达，一些细胞中与活性状态相关的表观遗传标记会增多（H3K9ac、H3K4me3）。可能是因为其他部分冗余的表观遗传机制能保证维持一条失活 X 染色体。

PRC2 蛋白被募集到失活 X 染色体上似乎是依赖了 *Xist* RNA。由于 PRC2 只是短

暂的结合到失活 X 染色体上，所以此复合物可能只在建立维持沉默的表观遗传标记（即 H3K27me3）时需要。目前还不知道 PRC1 是否直接识别这些标记，是否为失活 X 染色体的永久沉默所需要，但是已经发现 PRC1 组分与失活 X 染色体相关联。然而，DNA 甲基化已知伴随维持期存在，并且是失活 X 染色体的永久沉默所必需的。

PRC2 特异性参与了胚胎和胚胎外组织中 X 染色体单等位基因表达的调节，前者 X 染色体失活是随机的，后者父源 X 染色体几乎总是失活的（印记 X 染色体失活）。另外，最近还发现 PRC2 参与了一些印记常染色体的调控。例如，对六个不相关印记基因簇的 14 个印记位点分析表明，其中四个在 *eed* 突变小鼠中得到了双等位基因表达（Mager et al. 2003；详见第 19 章）。有趣的是所有失去印记表达的位点如果从父源继承一般都被抑制，而母源抑制的位点没有任何受到影响。最近发现 Ezh2 直接与哺乳类 DNA 甲基转移酶作用并与其活性相关（Viré et al. 2006），所以有可能 PRC2 通过 DNA 甲基化调控这些印记基因（见第 19 章）。

拟南芥的研究中也报道 PRC2 参与了印记基因表达的调控，其中父系等位基因的 *PHE1* 位点表达量大大提高（Köhler et al. 2005）。在 *E*（*z*）同源物 *MEA* 发生突变的植株中，母源 *PHE1* 等位基因特异性去抑制。MEA 也类似地调控自身的基因印记表达：早在生殖发育过程中，*mea* 突变体的母源 *MEA* 等位基因受到强烈的去抑制作用，不过这个作用并不需要 FIS 复合体其他组分（Baroux et al. 2006）。相反，在后续发育中，FIS 复合体能保证父源 *MEA* 等位基因的稳定抑制（Baroux et al. 2006；Gehring et al. 2006；Jullien et al. 2006）。后面的情况中，FIS 复合体参与了父源抑制的印记基因，这与哺乳类的情况类似。另外，如上所述，*MEA* 还保持母源 *PHE1*、*MEA* 等位基因低量表达（图 11-4）。

因为 PRC2 组分在植物、无脊椎动物以及哺乳动物中都存在，所以可以代表 1 个古老的分子构件，适合作为基因抑制的分子，基因抑制是在多细胞进化之前动植物单细胞祖先中就已经存在了的。这些例子也就说明了在动植物这两个进化有基因组印记的系中，PRC2 独立地被募集来调节其中印记基因的表达（Grossniklaus 2005）。

## 4.2 异常转录激活的后果

Bmi1 的错误调节会导致小鼠恶性淋巴瘤，那么人类 BMI1（PRC1 的组分之一）本身是否以相似的方式作用于癌症发展过程呢？越来越多的证据发现人类很多恶性淋巴瘤中都伴随有 PcG 基因表达改变（Raaphorst 2005）。例如，BMI1 在 B-细胞中过表达的量与肿瘤恶性程度有关，那么 PRC1 组分肯定在人类癌症发展过程中起到作用。然而，BMI1 在人类细胞中的靶基因似乎与小鼠淋巴细胞中的不同，因为在人类细胞中原癌基因过表达并没有伴随明显的 $p16^{INK4a}$ 下调。

PcG 基因过表达不仅在血液恶性肿瘤中观察到，还在实体肿瘤，包括髓母细胞瘤以及源于肝脏、结肠、乳腺、肺、睾丸和前列腺的肿瘤（图 11-8）。高增殖性的肺癌早期阶段常常有 PRC2 标记物——Ezh2 的高表达，表明从 PRC2 起始到 PRC1 维持（图 11-5）的这个熟识的级联反应可能也伴随着肿瘤细胞系的发展。

有趣的是，拟南芥的 PRC2 组分还在细胞增殖控制中起到至关重要的作用。尽管植

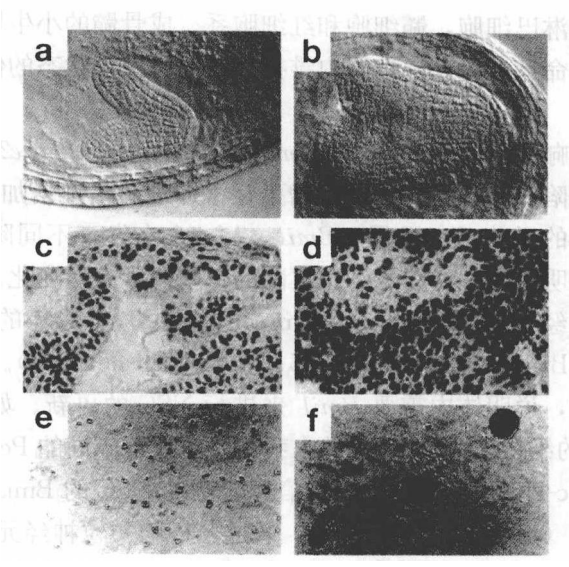



图 11-8　PRC2 调控哺乳动物和植物的细胞增殖

(a，b) 植物野生型和 *mea* 突变的卵细胞发育而来的胚。*MEA* 编码了 FIS 复合体的一个蛋白来调控细胞增殖。巨大的 *mea* 胚比处于相同发育阶段（晚心期）的野生型胚大很多。突变胚发育更慢，其层细胞的数量大约是野生型的两倍。(c，d) 正常和癌变的前列腺表皮。癌变表皮的 Ezh2 表达大大增加（抗 Ezh2 抗体标记），因此植物中 E(Z) 功能的丧失和人类 E(Z) 功能的过表达都可以导致细胞增殖的缺陷。(e，f) 对照大鼠、RING1 过表达大鼠的 1a 成纤维细胞。RING1 过表达导致其在软琼脂中出现不贴壁生长，这是癌转化细胞的典型特征。（a，b J.-P. Vielle-Calzada and U. Grossniklaus；c，d 经许可转印自 Kuzmichev et al. 2005；e，f 经许可转印自 Satijn and Otte1999）。

物异常生长不会导致癌症或死亡，但仍然需要严格控制细胞增殖来维持正常发育。*fis* 类基因发生突变的植株，开花植物的两个受精产物，胚和胚乳过度增殖，导致种子败育（图 11-8）（Grossniklaus et al. 2001；Hsieh et al. 2003；Guitton and Berger 2005）。*clf* 和 *swn* 这两个植物 *E(z)* 同源物的双突变体中也观察到了细胞增殖的效果。这些植物在萌芽后先进行了正常的种子发育，但是之后产生了一堆增殖未分化的组织（愈伤组织），而不是叶（Chanvivattana et al. 2004）。

尽管还不知道 PRC2 是怎么控制植物细胞增殖的，但这一过程很可能涉及与 RBR1——植物的视网膜母细胞瘤（retinoblastoma，Rb）蛋白同源物的相互作用（Ebel et al. 2004；Mosquna et al. 2004）。*FIS* 类基因的突变体不仅在受精后的种子发育中显示出扩增缺陷，还能在未受精的胚乳中阻止增殖。后面的这种表型是 *rbr1* 突变体也有的，这就与 Rb 通路相联系了。值得注意的是哺乳动物 Rb 通路和 PRC2 之间也有联系（Bracken et al. 2003），表明动植物之间有保守的调节网络。

## 4.3　维持干细胞命运

干细胞在医疗中起到越来越重要的作用。它们可以为受损组织的复愈提供前体，因此是再生医疗的百宝箱。我们最了解的是血细胞系中干细胞的身份和地位，因此这个百宝箱位于血细胞系也不足为怪了。

造血干细胞（hematopoietic stem cell，HSC）通过自身更新来维持血细胞库，同时

产生子细胞，分化成淋巴细胞，髓细胞和红细胞系。成骨髓的小生境为干细胞提供特异的外界信号以维持其命运。另外，细胞内维持“干细胞性”状态的信息来源可能依赖于PcG系统。

*PRC1*基因受影响的小鼠突变体（如*bmi1/mel18*，*mph1/rae28*及*m33*；表11-1）其造血系统有多种缺陷，如脾、胸腺中的增生症（即细胞扩增增加），B、T细胞减少，淋巴前体对细胞因子的增殖反应受损。*Bmi1*、*Mel18*在发育不同阶段干细胞自身更新中都是需要的，这说明胚胎和成体干细胞之间的靶基因库有所变化。

PcG系统还是神经干细胞（neural stem cell，NSC）中必需的，*bmi1*小鼠突变体中就有神经元缺陷（Bruggeman et al. 2005；Zencak et al. 2005）。这些小鼠在出生后尤其缺失了脑神经元，说明体内需要*Bmi1*来进行NSC的更新。如造血系统的发现一样，似乎胚胎NSC的维持、成体NSC的自我更新也有着不同的PcG网络调控方式。

外界信号如sonic-Hedgehog蛋白级联信号能调控NSC的Bmi1反应并且保证其增殖或自我更新的能力（Leung et al. 2004）。通过分析脑颗粒神经元前体细胞（cerebellar granule neuron progenitors，CGNP）的发育确定了一些控制PcG抑制的外部信号。出生后的增殖波是由Purkinje细胞分泌的信号因子Sonichedgehog蛋白（Shh）诱导产生的。Shh信号分支能控制N-Myc和Bmi1水平（图11-9）。因此，Bmi1缺陷型CGNP在Shh激发时的增殖反应也有缺陷。Shh信号可以最终通过调控下游Rb通路（通过N-

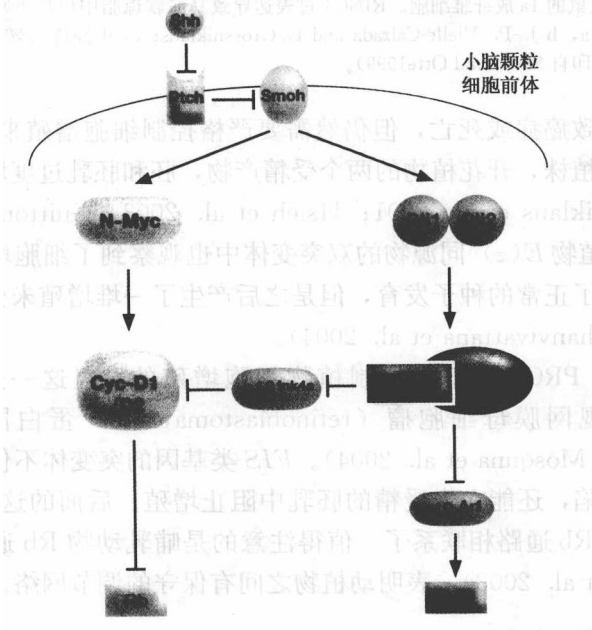



图11-9 Shh信号维持脑前体细胞的增殖/自我更新

Shh信号级联作用调控了Rb通路和p53通路，通过Bmi1控制*p16/p19*扩增检验点。Smoh被Shh调控子Ptch抑制，使核仁得到下游信号。信号的一部分诱导N-myc、Cyclin D1、D2表达，另一部分通过Gli效应器激活Bmi1（选自ValkLingbeek et al. 2004）。

myc 和 Bmi1/p16$^{INK4a}$）和 p53（通过 Bmi1/p19$^{ARF}$）通路来控制这些干细胞的扩增。这个机制解释了为何 Shh 信号的超激活导致髓母细胞瘤的发展（Leung et al. 2004）。HSC 受到相似的 ihh 信号（Indian hedgehog）控制途径的调控。NSC 中 *Hoxd8*、*Hoxd9*、*Hoxc9* 位点的表达受 Bmi1 的控制。合适的 *HOX* 表达方式决定了必要的干细胞命运。

事实上，由于干细胞的细胞命运是确定的，所以 PcG 系统在有丝分裂中保持这种特定命运也不足为奇。未来有意义的研究是找出不同干细胞群的 PcG 系统目标库，了解如何干预这种维持系统以人为控制干细胞命运的重编程。现在还不清楚 PcG 是否在植物的干细胞维持中能够起到作用，植物细胞是多能的并且有在一定条件下形成一个新植株的潜能，能够确信在这些植物细胞的重编程过程中有 PcG 的调节。其实，缺少 *E*(*z*)同源物 *CLF* 和 *SWN* 的植株在萌芽后产生了一团未分化的细胞，也表明 PcG 基因在维持分化状态时是需要的（Chanvivattana et al. 2004）。

## 5. 总结及展望

自从一开始遗传学发现了在第二、三条腿上有多余性梳的果蝇突变体以来，我们对于 PcG 表观遗传调控的理解有着令人瞩目的发展。这又促进了一类新的调控因子的发现，它们为基础表观遗传过程（如植物春化、哺乳动物 X 染色体沉默）所需要。染色质结构及组蛋白的不同修饰形式也大大影响了遗传信息的控制。在发育决定之后，PcG 蛋白直接参与产生表观遗传标记（如 H3K27me3）。同一组蛋白通过 PRC1 蛋白的作用“读出”（即显示出高亲和性）这些表观遗传标记并将其翻译成一种稳定的转录抑制状态。在果蝇这种模式生物中，我们对 PcG 复合体锚定在 PRE 上的方式了解相对透彻一些，因为我们知道了其中哪一类靶基因受长程抑制。然而迄今为止，在其他生物体中没有发现 PRE。尽管植物和脊椎动物 PcG 蛋白的基本功能是一样的，但还不清楚二者基因组的哪一部分受到抑制，还有 PcG 蛋白如何被靶向其作用位点的。另外，我们需要搞清楚这些明显动态的蛋白组怎么通过化学平衡的方式产生一种稳定的转录抑制状态。

PcG 研究的其他问题主要集中于抑制状态如何被遗传，这也是表观遗传学的精髓所在。DNA 复制和有丝分裂中要传递基因表达的状态需要什么样的分子标记呢？我们知道 trxG 和 PcG 蛋白的联合作用可以使基因表达维持一种激活或沉默的状态，那么两种状态是否需要相应的表观遗传标记传递到子细胞，还是只需要一种标记，即可忽略其他的状态？在分裂间期，PcG 蛋白造成转录沉默的机制变得越来越清楚。未来研究的焦点在于关系到基因表达状态的信息如何通过 DNA 复制阶段而忠实地传递到有丝分裂的子细胞中。

朱慧君　译
王大亮　校
孙方霖　审校

# 参考文献


Akasaka T., van Lohuizen M., van der Lugt N., Mizutani-Koseki Y., Kanno M., Taniguchi M., Vidal M., Alkema M., Berns A., and Koseki H. 2001. Mice doubly deficient for the *Polycomb* Group genes *Mel18* and *Bmi1* reveal synergy and requirement for maintenance but not initiation of *HOX* gene expression. *Development* **128:** 1587–1597.

Ali J.Y. and Bender W. 2004. Cross-regulation among the *Polycomb* group genes in *Drosophila melanogaster*. *Mol. Cell. Biol.* **24:** 7737–7747.

Ausin I., Alonso-Blanco C., Jarillo J.A., Ruiz-Garcia L., and Martinez-Zapater J.M. 2004. Regulation of flowering time by FVE, a retinoblastoma-associated protein. *Nat. Genet.* **36:** 162–166.

Bannister A.J., Zegerman P., Partridge J.F., Miska E.A., Thomas J.O., Allshire R.C., and Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromodomain. *Nature* **410:** 120–124.

Baroux C., Gagliardini V., Page D., and Grossniklaus U. 2006. Dynamic regulatory interactions of *Polycomb* group genes: *MEDEA* autoregulation is required for imprinted gene expression in *Arabidopsis*. *Genes Dev.* **20:** 1081–1086.

Bastow R., Mylne J.S., Lister C., Lippman Z., Martienssen R.A., and Dean C. 2004. Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature* **427:** 164–167.

Birve A., Sengupta A.K., Beuchle D., Larsson J., Kennison J.A., Rasmuson-Lestander A., and Müller J. 2001. *Su(z)12*, a novel *Drosophila Polycomb* group gene that is conserved in vertebrates and plants. *Development* **128:** 3371–3379.

Bracken A.P., Pasini D., Capra M., Prosperini E., Colli E., and Helin K. 2003. *EZH2* is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. *EMBO J.* **22:** 5323–5335.

Breiling A., Bonte E., Ferrari S., Becker P.B., and Paro R. 1999. The *Drosophila Polycomb* protein interacts with nucleosomal core particles in vitro via its repression domain. *Mol. Cell. Biol.* **19:** 8451–8460.

Bruggeman S.W.M., Valk-Lingbeek M.E., van der Stoop P.P.M., Jacobs J.J.L., Kieboom K., Tanger E., Hulsman D., Leung C., Arsenijevic Y., Marino S., and van Lohuizen M. 2005. *Ink4a* and *Arf* differentially affect cell proliferation and neural stem cell self-renewal in *Bmi1*-deficient mice. *Genes Dev.* **19:** 1438–1443.

Cao R., Tsukada Y., and Zhang Y. 2005. Role of Bmi-1 and Ring1A in H2A ubiquitylation and *HOX* gene silencing. *Mol. Cell* **20:** 845–854.

Cao R., Wang L.J., Wang H.B., Xia L., Erdjument-Bromage H., Tempst P., Jones R.S., and Zhang Y. 2002. Role of histone H3 lysine 27 methylation in *Polycomb*-group silencing. *Science* **298:** 1039–1043.

Carrington E.A. and Jones R.S. 1996. The *Drosophila Enhancer of zeste* gene encodes a chromosomal protein: Examination of wild-type and mutant protein distribution. *Development* **122:** 4073–4083.

Chanvivattana Y., Bishopp A., Schubert D., Stock C., Moon Y.H., Sung Z.R., and Goodrich J. 2004. Interaction of *Polycomb*-group proteins controlling flowering in *Arabidopsis*. *Development* **131:** 5263–5276.

Core N., Charroux B., McCormick A., Vola C., Fasano L., Scott M.P., and Kerridge S. 1997. Transcriptional regulation of the *Drosophila* homeotic gene *teashirt* by the homeodomain protein Fushi tarazu. *Mech. Dev.* **68:** 157–172.

Czermin B., Melfi R., McCabe D., Seitz V., Imhof A., and Pirrotta V. 2002. *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal *Polycomb* sites. *Cell* **111:** 185–196.

Dejardin J., Rappailles A., Cuvier O., Grimaud C., Decoville M., Locker D., and Cavalli G. 2005. Recruitment of *Drosophila Polycomb* group proteins to chromatin by DSP1. *Nature* **434:** 533–538.

Dellino G.I., Schwartz Y.B., Farkas G., McCabe D., Elgin S.C., and Pirrotta V. 2004. *Polycomb* silencing blocks transcription initiation. *Mol. Cell* **13:** 887–893.

del Mar Lorente D., Marcos-Gutierrez C., Perez C., Schoorlemmer J., Ramirez A., Magin T., and Vidal M. 2000. Loss- and gain-of-function mutations show a *Polycomb* group function for Ring1A in mice. *Development* **127:** 5093–5100.

de Napoles M., Mermoud J.E., Wakao R., Tang Y.A., Endoh M., Appanah R., Nesterova T.B., Silva J., Otte A.P., Vidal M., et al. 2004. *Polycomb* group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7:** 663–676.

Ebel C., Mariconti L., and Gruissem W. 2004. Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte. *Nature* **429:** 776–780.

Fang J., Chen T.P., Chadwick B., Li E., and Zhang Y. 2004. Ring1b-mediated H2A ubiquitination associates with inactive X chromosomes and is involved in initiation of X inactivation. *J. Biol. Chem.* **279:** 52812–52815.

Ficz G., Heintzmann R., and Arndt Jovin D.J. 2005. *Polycomb* group protein complexes exchange rapidly in living *Drosophila*. *Development* **132:** 3963–3976.

Fischle W., Wang Y., Jacobs S.A., Kim Y., Allis C.D., and Khorasanizadeh S. 2003. Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes Dev.* **17:** 1870–1881.

Fong Y., Bender L., Wang W., and Strome S. 2002. Regulation of the different chromatin states of autosomes and X chromosomes in the germ line of *C. elegans*. *Science* **296:** 2235–2238.

Francis N.J., Kingston R.E., and Woodcock C.L. 2004. Chromatin compaction by a *Polycomb* group protein complex. *Science* **306:** 1574–1577.

Francis N.J., Saurin A.J., Shao Z., and Kingston R.E. 2001. Reconstitution of a functional core *Polycomb* repressive complex. *Mol. Cell* **8:** 545–556.

Gehring M., Huh J.H., Hsieh T.F., Penterman J., Choi Y., Harada J.J., Goldberg R.B., and Fischer R.L. 2006. DEMETER DNA glycosylase establishes *MEDEA Polycomb* gene self-imprinting by allele-specific demethylation. *Cell* **124:** 495–506.

Gendall A.R., Levy Y.Y., Wilson A., and Dean C. 2001. The *VERNALIZATION2* gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. *Cell* **107:** 525–535.

Goodrich J., Puangsomlee P., Martin M., Long D., Meyerowitz E.M., and Coupland G. 1997. A *Polycomb*-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* **386:** 44–51.

Grossniklaus U. 2005. Genomic imprinting in plants: A predominantly maternal affair. In *Annual plant reviews: Plant epigenetics* (ed. P. Meyer), pp. 174–200. Blackwell, Sheffield, United Kingdom.

Grossniklaus U., Spillane C., Page D.R., and Köhler C. 2001. Genomic imprinting and seed development: Endosperm formation with and without sex. *Curr. Opin. Plant Biol.* **4:** 21–27.

Grossniklaus U., Vielle-Calzada J.P., Hoeppner M.A., and Gagliano W.B. 1998. Maternal control of embryogenesis by *MEDEA*, a *Polycomb* group gene in *Arabidopsis*. *Science* **280:** 446–450.

Guitton A.E. and Berger F. 2005. Control of reproduction by *Polycomb* Group complexes in animals and plants. *Int. J. Dev. Biol.* **49:** 707–716.

Hackett W.P., Cordero R.E., and Sinivasan C. 1987. Apical meristem characteristics and activity in relation to juvenility in *Hedera*. In *Manipulation of flowering* (ed. J.G. Atherton), pp. 93–99. Butterworth, London.

Hadorn E. 1968. Transdetermination in cells. *Sci. Am.* **219:** 110.

Heard E. 2004. Recent advances in X-chromosome inactivation. *Curr. Opin. Cell Biol.* **16:** 247–255.

Hennig L., Bouveret R., and Gruissem W. 2005. MSI1-like proteins: An escort service for chromatin assembly and remodeling complexes. *Trends Cell Biol.* **15:** 295–302.

Hennig L., Taranto P., Walser M., Schonrock N., and Gruissem W. 2003.

*Arabidopsis* MSI1 is required for epigenetic maintenance of reproductive development. *Development* **130:** 2555–2565.

Hsieh T.F., Hakim O., Ohad N., and Fischer R.L. 2003. From flour to flower: How *Polycomb* group proteins influence multiple aspects of plant development. *Trends Plant Sci.* **8:** 439–445.

Jacobs J.J.L., Scheijen B., Voncken J.W., Kieboom K., Berns A., and van Lohuizen M. 1999. Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. *Genes Dev.* **13:** 2678–2690.

Jullien P.E., Katz A., Oliva M., Ohad N., and Berger F. 2006. *Polycomb* group complexes self-regulate imprinting of the *Polycomb* group gene *MEDEA* in *Arabidopsis*. *Curr. Biol.* **16:** 486–492.

Kagey M.H., Melhuish T.A., and Wotton D. 2003. The *Polycomb* protein Pc2 is a SUMO E3. *Cell* **113:** 127–137.

Kennison J.A. 1995. The *Polycomb* and *trithorax* group proteins of *Drosophila*: Trans-regulators of homeotic gene function. *Annu. Rev. Genet.* **29:** 289–303.

Kim H.J., Hyun Y., Park J.Y., Park M.J., Park M.K., Kim M.D., Kim H.J., Lee M.H., Moon J., Lee I., and Kim J. 2004. A genetic link between cold responses and flowering time through *FVE* in *Arabidopsis thaliana*. *Nat. Genet.* **36:** 167–171.

Kinoshita T., Harada J.J., Goldberg R.B., and Fischer R.L. 2001. *Polycomb* repression of flowering during early plant development. *Proc. Natl. Acad. Sci.* **98:** 14156–14161.

Klebes A., Sustar A., Kechris K., Li H., Schubiger G., and Kornberg T.B. 2005. Regulation of cellular plasticity in *Drosophila* imaginal disc cells by the *Polycomb* group, *trithorax* group and *lama* genes. *Development* **132:** 3753–3765.

Klymenko T. and Muller J. 2004. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by *Polycomb* group proteins. *EMBO Rep.* **5:** 373–377.

Köhler C., Page D.R., Gagliardini V., and Grossniklaus U. 2005. The *Arabidopsis thaliana* MEDEA *Polycomb* group protein controls expression of *PHERES1* by parental imprinting. *Nat. Genet.* **37:** 28–30.

Köhler C., Hennig L., Bouveret R., Gheyselinck J., Grossniklaus U., and Gruissem W. 2003a. *Arabidopsis* MSI1 is a component of the MEA/FIE *Polycomb* group complex and required for seed development. *EMBO J.* **22:** 4804–4814.

Köhler C., Hennig L., Spillane C., Pien S., Gruissem W., and Grossniklaus U. 2003b. The *Polycomb* group protein MEDEA regulates seed development by controlling expression of the MADS-box gene *PHERES1*. *Genes Dev.* **17:** 1540–1553.

Kuzmichev A., Jenuwein T., Tempst P., and Reinberg D. 2004. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol. Cell* **14:** 183–193.

Kuzmichev A., Nishioka K., Erdjument-Bromage H., Tempst P., and Reinberg D. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16:** 2893–2905.

Kuzmichev A., Margueron R., Vaquero A., Preissner T.S., Scher M., Kirmizis A., Ouyang X., Brockdorff N., Abate Shen C., Farnham P., and Reinberg D. 2005. Composition and histone substrates of *Polycomb* repressive group complexes change during cellular differentiation. *Proc. Natl. Acad. Sci.* **102:** 1859–1864.

Lavigne M., Francis N.J., King I.F., and Kingston R.E. 2004. Propagation of silencing; recruitment and repression of naive chromatin in *trans* by *Polycomb* repressed chromatin. *Mol. Cell* **13:** 415–425.

Lee N., Maurange C., Ringrose L., and Paro R. 2005. Suppression of *Polycomb* group proteins by JNK signalling induces transdetermination in *Drosophila* imaginal discs. *Nature* **438:** 234–237.

Leung C., Lingbeek M., Shakhova O., Liu J., Tanger E., Saremaslani P., van Lohuizen M., and Marino S. 2004. Bmi1 is essential for cerebellar development and is overexpressed in human medulloblastomas. *Nature* **428:** 337–341.

Levine S.S., King I.F., and Kingston R.E. 2004. Division of labor in *Polycomb* group repression. *Trends Biochem. Sci.* **29:** 478–485.

Levine S.S., Weiss A., Erdjument Bromage H., Shao Z., Tempst P., and Kingston R.E. 2002. The core of the *Polycomb* Repressive Complex is compositionally and functionally conserved in flies and humans. *Mol. Cell. Biol.* **22:** 6070–6078.

Lewis E.B. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276:** 565–570.

Luo M., Bilodeau P., Koltunow A., Dennis E.S., Peacock W.J., and Chaudhury A.M. 1999. Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* **96:** 296–301.

Mager J., Montgomery N.D., de Villena F.P.M., and Magnuson T. 2003. Genome imprinting regulated by the mouse *Polycomb* group protein Eed. *Nat. Genet.* **33:** 502–507.

Marx J. 2005. Developmental biology—Combing over the *Polycomb* group proteins. *Science* **308:** 624–626.

Moon Y.H., Chen L., Pan R.L., Chang H.S., Zhu T., Maffeo D.M., and Sung Z.R. 2003. EMF genes maintain vegetative development by repressing the flower program in *Arabidopsis*. *Plant Cell* **15:** 681–693.

Mosquna A., Katz A., Shochat S., Grafi G., and Ohad N. 2004. Interaction of FIE, a *Polycomb* protein, with pRb: a possible mechanism regulating endosperm development. *Mol. Genet. Genomics* **271:** 651–657.

Müller J., Hart C.M., Francis N.J., Vargas M.L., Sengupta A., Wild B., Miller E.L., O'Connor M.B., Kingston R.E., and Simon J.A. 2002. Histone methyltransferase activity of a *Drosophila* polycomb group repressor complex. *Cell* **111:** 197–208.

Ohad N., Yadegari R., Margossian L., Hannon M., Michaeli D., Harada J.J., Goldberg R.B., and Fischer R.L. 1999. Mutations in *FIE*, a WD *Polycomb* group gene, allow endosperm development without fertilization. *Plant Cell* **11:** 407–415.

Otte A.P. and Kwaks T.H. 2003. Gene repression by *Polycomb* group protein complexes: A distinct complex for every occasion? *Curr. Opin. Genet. Dev.* **13:** 448–454.

Page D.R. and Grossniklaus U. 2002. The art and design of genetic screens: *Arabidopsis thaliana*. *Nat. Rev. Genet.* **3:** 124–136.

Paro R. and Hogness D.S. 1991. The *Polycomb* protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl. Acad. Sci.* **88:** 263–267.

Poux S., McCabe D., and Pirrotta V. 2001. Recruitment of components of *Polycomb* group chromatin complexes in *Drosophila*. *Development* **128:** 75–85.

Raaphorst F.M. 2005. Deregulated expression of *Polycomb*-group oncogenes in human malignant lymphomas and epithelial tumors. *Hum. Mol. Genet.* **14:** R93–100.

Rea S., Eisenhaber F., O'Carroll N., Strahl B.D., Sun Z.W., Schmid M., Opravil S., Mechtler K., Ponting C.P., Allis C.D., and Jenuwein T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406:** 593–599.

Reyes J.C. and Grossniklaus U. 2003. Diverse functions of *Polycomb* group proteins during plant development. *Semin. Cell Dev. Biol.* **14:** 77–84.

Ringrose L., Ehret H., and Paro R. 2004. Distinct contributions of histone H3 lysine 9 and 27 methylation to locus-specific stability of *Polycomb* complexes. *Mol. Cell* **16:** 641–653.

Ringrose L., Rehmsmeier M., Dura J.M., and Paro R. 2003. Genome-wide prediction of Polycomb/Trithorax response elements in *Drosophila melanogaster*. *Dev. Cell* **5:** 759–771.

Sanchez-Elsner T., Gou D., Kremmer E., and Sauer F. 2006. Noncoding RNAs of trithorax response elements recruit *Drosophila* Ash1 to *Ultrabithorax*. *Science* **311:** 1118–1123.

Satijn D.P. and Otte A.P. 1999. RING1 interacts with multiple *Polycomb*-group proteins and displays tumorigenic activity. *Mol. Cell. Biol.* **19:** 57–68.

Saurin A.J., Shiels C., Williamson J., Satijn D.P.E., Otte A.P., Sheer D., and Freemont P.S. 1998. The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. *J. Cell Biol.* **142:** 887–898.

Schmitt S., Prestel M., and Paro R. 2005. Intergenic transcription through a *Polycomb* group response element counteracts silencing. *Genes Dev.* **19:** 697–708.

Sung S. and Amasino R.M. 2004a. Vernalization and epigenetics: How plants remember winter. *Curr. Opin. Plant Biol.* **7:** 4–10.

———. 2004b. Vernalization in *Arabidopsis thaliana* is mediated by

the PHD finger protein VIN3. *Nature* **427:** 159–164.

Tie F., Furuyama T., Prasad-Sinha J., Jane E., and Harte P.J. 2001. The *Drosophila Polycomb* group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. *Development* **128:** 275–286.

Tschiersch B., Hofmann A., Krauss V., Dorn R., Korge G., and Reuter G. 1994. The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* **13:** 3822–3831.

Valk-Lingbeek M.E., Bruggeman S.W.M., and van Lohuizen M. 2004. Stem cells and cancer: The *Polycomb* connection. *Cell* **118:** 409–418.

van der Lugt N.M., Domen J., Linders K., van Roon M., Robanus-Maandag E., te Riele H., van der Valk M., Deschamps J., Sofroniew M., van Lohuizen M., et al. 1994. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. *Genes Dev.* **8:** 757–769.

van Lohuizen M., Verbeek S., Scheijen B., Wientjens E., van der Gulden H., and Berns A. 1991. Identification of cooperating oncogenes in Eμ-myc transgenic mice by provirus tagging. *Cell* **65:** 737–752.

Viré E., Brenner C., Deplus R., Blanchon L., Fraga M., Didelot C., Morey L., van Eynde A., Bernhard D., Vanderwinden J.M., et al. 2006. The *Polycomb* group protein EZH2 directly controls DNA methylation. *Nature* **439:** 871–874.

Wang L., Brown J.L., Cao R., Zhang Y., Kassis J.A., and Jones R.S. 2004. Hierarchical recruitment of *Polycomb* group silencing complexes. *Mol. Cell* **14:** 637–646.

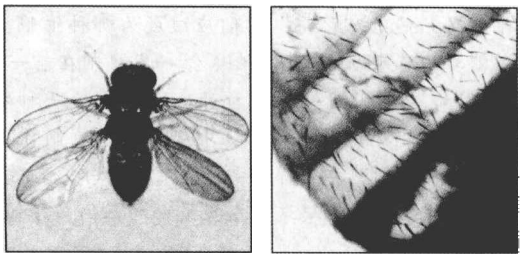
Yamamoto Y., Girard F., Bello B., Affolter M., and Gehring W.J. 1997. The *cramped* gene of *Drosophila* is a member of the *Polycomb*-group, and interacts with *mus209*, the gene encoding proliferating cell nuclear antigen. *Development* **124:** 3385–3394.

Yamamoto K., Sonoda M., Inokuchi J., Shirasawa S., and Sasazuki T. 2004. *Polycomb* group Suppressor of zeste 12 links Heterochromatin Protein 1α and Enhancer of zeste 2. *J. Biol. Chem.* **279:** 401–406.

Yoshida N., Yanai Y., Chen L.J., Kato Y., Hiratsuka J., Miwa T., Sung Z.R., and Takahashi S. 2001. EMBRYONIC FLOWER2, a novel *Polycomb* group protein homolog, mediates shoot development and flowering in *Arabidopsis*. *Plant Cell* **13:** 2471–2481.

Zencak D., Lingbeek M., Kostic C., Tekaya M., Tanger E., Hornfeld D., Jaquet M., Munier F.L., Schorderet D.F., van Lohuizen M., and Arsenijevic Y. 2005. Bmi1 loss produces an increase in astroglial cells and a decrease in neural stem cell population and proliferation. *J. Neurosci.* **25:** 5774–5783.

# 第 12 章



# 三胸蛋白家族（trxG）与转录调控

Robert E. Kingston[1] and John W. Tamkun[2]

[1]*Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114*

[2]*Department of Molecular, Cell and Developmental Biology, University of California, Santa Cruz, California 95064*

## 本章目录

1. 前言
   - 1.1 鉴定维持细胞特定状态所需基因
   - 1.2 其他物种中 trxG 蛋白
   - 1.3 真核转录中 trxG 蛋白具有的多样化功能
2. trxG 蛋白与染色质
   - 2.1 trxG 蛋白在 ATP 依赖的染色质重塑中起作用
   - 2.2 trxG 蛋白共价修饰核小体组蛋白
3. trxG 蛋白与通用转录装置
4. 其他 trxG 蛋白的生化功能
5. trxG 蛋白间功能性相互作用
6. trxG 蛋白：激活子或是反抑制子
7. 结论和展望

参考文献

## 概要

一个生物体内的所有细胞必须能够“记得”它们应该成为哪种细胞类型，这个叫做“细胞记忆”或“转录记忆”的过程需要两类基础的调控机制。第一类机制在前一章已经讨论过，是要维持一些基因的“关闭”状态，因为这些基因一旦被打开，则会导致一种特异的非正常细胞类型。Polycomb 家族（PcG）蛋白主要功能就是在细胞记忆中起到这种抑制作用。第二类则是维持关键基因处于“开启”状态的机制。任何细胞类型都需要表达主控蛋白，这些蛋白提供了此细胞类型所必需的特异性功能。生物体的一生当中，编码这些主控蛋白的基因必须维持在一种“开启”状态，才能在体内维持为适当的细胞类型。

左题图是一个多重翅果蝇，这种惊人的现象表明，当一个主控蛋白基因没能维持表达时能够导致表型发生剧烈的变化。维持基因处于“开启”状态的蛋白叫做 Trithorax 家族（trxG），这得名于这群调控蛋白中奠基性的成员——*trithorax* 基因。trxG 家族包括一大群蛋白，分别具有各种各样的功能。在表观遗传维持“开启”状态的机制中，这些蛋白所起的作用在这一点上看比 PcG 蛋白的抑制作用显得更复杂。第一个复杂性在于，从任何基因上活跃地转录 RNA 都需要大量蛋白和多种机制，而抑制作用可以用的机制相对简单，阻断所有蛋白接近基因就可以做到；相反，基因的激活要求有很多步骤，其中每一步都可能对维持基因“开启”状态起到作用，因此 trxG 蛋白工作的时期可能也很多。

考虑到 trxG 蛋白的性质，第二个复杂性在于它们不但可以起到激活作用，还可以在其他条件下起到抑制性作用。这看似费解，但是根据基因的精细结构，相同蛋白的功能可能在一种情况下帮助某个基因激活，另一种情况下却使另一个基因受抑制。这时的 trxG 蛋白就不是专门维持基因表达了，而是可以在细胞内起到多种作用。也正是这些复杂之处引发了几个尚未解决的有趣问题——为什么激活转录所需的蛋白中有一些对于维持转录也是重要的呢？只有这些蛋白具有维持转录的特定功能吗？还是一些蛋白仅仅由于进化的偶然性成为调控发育中关键基因的重要因子而在维持转录中必需？

下面即将提到，一些 trxG 蛋白调控染色质结构使用了与 PcG 蛋白相反的机制，trxG 蛋白能够给染色质加上共价修饰，或者可以通过改变染色质构建元件——核小体的结构或位置来改变染色质。其他 trxG 蛋白是转录装置的一部分。因此相对于 PcG 蛋白，更多类型复合物中都含有 trxG 蛋白，它们也很可能在表观遗传机制中起到更复杂的作用。

## 1. 前　言

很多发育的决定过程，包括细胞命运决定都与早期胚胎的位置信息相对应。这些决定依赖于基因表达的变化，使拥有相同遗传蓝本的细胞得到独特的身份并沿着不同的途径进行发育。奠定细胞命运决定基础的基因表达变化是可以遗传的，细胞命运一旦决定，即使细胞分裂了很多代，发育了很长时间也很少变化。了解细胞确定状态是如何维持的是发育生物学家和分子生物学家长期以来的一个目标。

在研究果蝇同源异型（*Hox*）基因时，人们发现了很多调控蛋白与维持可遗传的基因表达状态相关。*Hox* 基因编码同源盒转录因子，这些因子可以调控下游一组目标基因的转录，这反过来又决定了体节的特异性标志（Gellon and McGinnis 1998）。果蝇的 *Hox* 基因存在于两个基因家族复合物中：即触角足-复合物（Antennapedia complex，ANT-C）及双胸-复合物（Bithorax，BX-C）。其中前者包括 *Hox* 家族之 *lab*、*Dfd*、*Scr* 和 *Antp* 基因；后者包括 *Hox* 家族之 *Ubx*，*abdA* 以及 *AbdB* 基因（Duncan 1987；Kaufman et al. 1990）。每个 *Hox* 基因都决定了成为发育中果蝇一个特定体节或前后轴沿线体节组。例如，*Antp* 决定了形成第二胸节，包括翅膀后的平衡器官。因此 *Hox* 基因编码的转录因子就起到主控开关的功能，可以直接转向发育中的可选途径。

*Hox* 基因的转录必须得到精确的调控，因为如果它们不适当表达会导致细胞命运剧烈的改变（Simon 1995；Simon and Tamkun 2002）。如头节 *Antp* 受抑制会使触角转化为腿，而胸节 *Ubx* 失活则导致平衡器官成为翅膀。果蝇 Hox 转录的初始状态是在胚胎形成早期由分节基因编码的转录因子建立的。裂缺基因（g p gene）、配对规则基因（pair-rule gene）以及体节极性基因（segment polarity gene）等分节基因编码的蛋白，将早期胚胎分为 14 个相同的体节，这些蛋白还建立起 *Hox* 转录的初始状态，这也是体节向具有不同标识、形态阶段发育的第一步。受体节限制的 *Hox* 转录形式建立起来之后就必须在随后的胚胎期、幼虫期和成蛹期一直维持下去，这样才能维持各体节的标志。由于大多数分节基因只在早期发育中短暂表达，因此这项维持功能由另外两组调控蛋白完成——抑制蛋白 PcG 以及转录调控子 TrxG（图 12-1）。这样，*Hox* 的转录调控至少包含（由分节基因）建立和（由 PcG、TrxG 基因）维持两个不同阶段（图 12-2）。

包括 *trx*、*ash*1、*ash*2 和 *fsh* 在内的 trxG 成员最初在与 *Hox* 功能失活突变（表型）相似的突变中发现（图 12-3）（Kennison 1995）。例如，*trx*（trxG 家族奠基性成员）突变引起平衡棒部分形成翅膀（*Ubx* 转录减少造成）；前足转化为中足（*Scr* 转录减少造成）；以及后腹节具有更多前部特征（*abdA* 和 *AbdB* 转录减少造成）。其他大量 trxG 成员是在筛选 *Pc* ［*Su*（*Pc*）］突变的基因外抑制子时发现的（Kennison and Tamkun 1988）。这些遗传筛选背后的推理是这样的：PcG 抑制子表达水平减少了，则作为补偿，维持活性状态的蛋白水平应当下降（图 12-4）。*brm* 和大量其他 *Su*（*Pc*）位点都是由此方法找到的，这也使 trxG 成员的总数超过了 16 个（表 12-1）。许多其他蛋白也被划分为 trxG 家族，这是由于标准不严格产生的，这些标准包括与已知 trxG 蛋白序列同源，与 trxG 蛋白有物理上的联系，在体内外对 Hox 转录的效应和生化活性等。

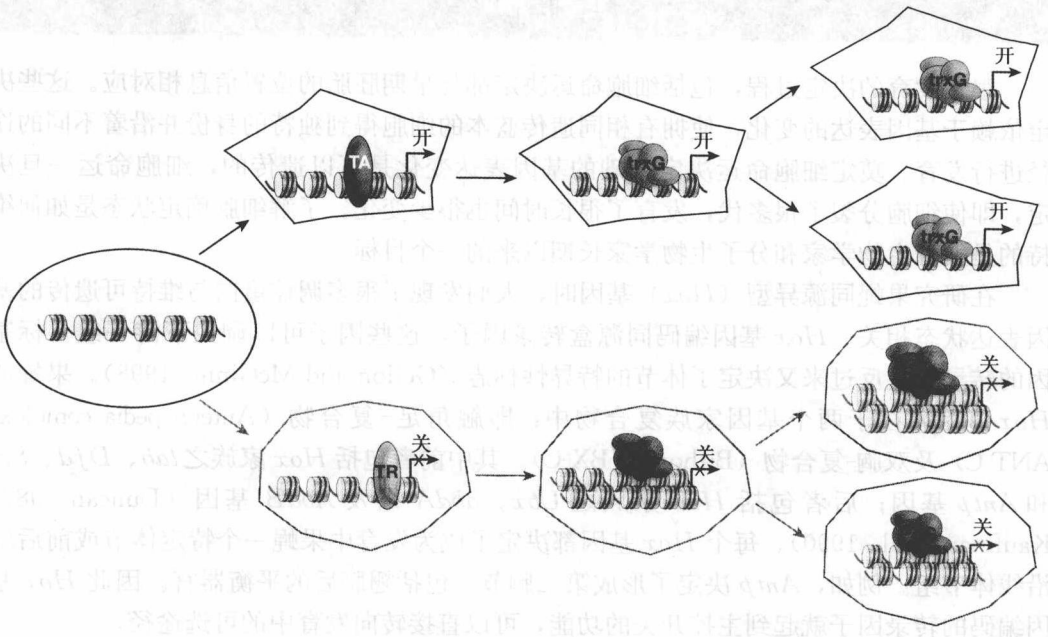



图 12-1 细胞记忆的概念

本图重点描述了 trxG 复合体对可继承的活性基因表达状态的维持作用，也显示了 PcG 复合体的相反作用——它们维持了果蝇 Hox 基因簇初始的沉默状态。

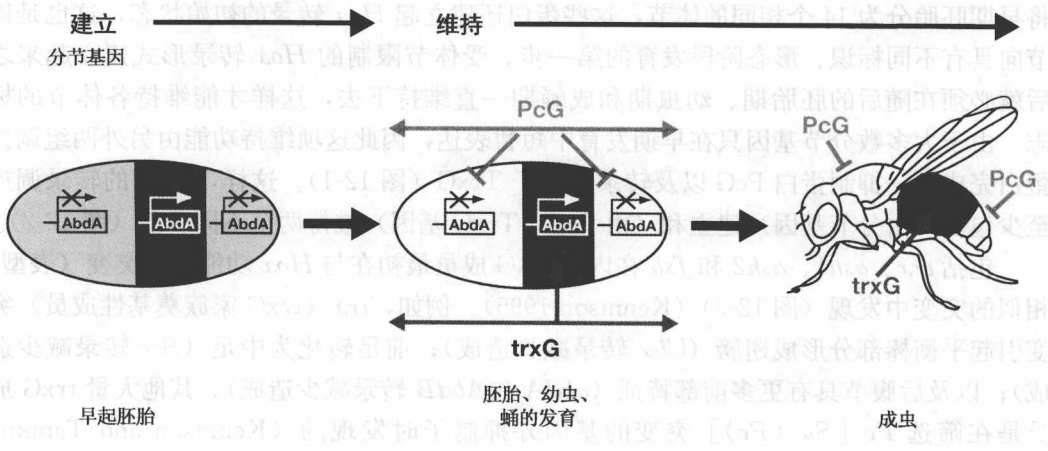



图 12-2 *Hox* 的转录调控

*abdA* 转录和其他 *Hox* 基因的界限是由分节蛋白建立的。这些分节蛋白包括了裂缺基因、配对规则基因的产物，它们将胚胎分成 14 个相同的体节。随后的发育过程中，*Hox* 转录的“关闭”或“开启”状态就由广泛表达的 trxG 激活子成员以及 PcG 抑制子成员来维持，具体机制现在还不清楚。

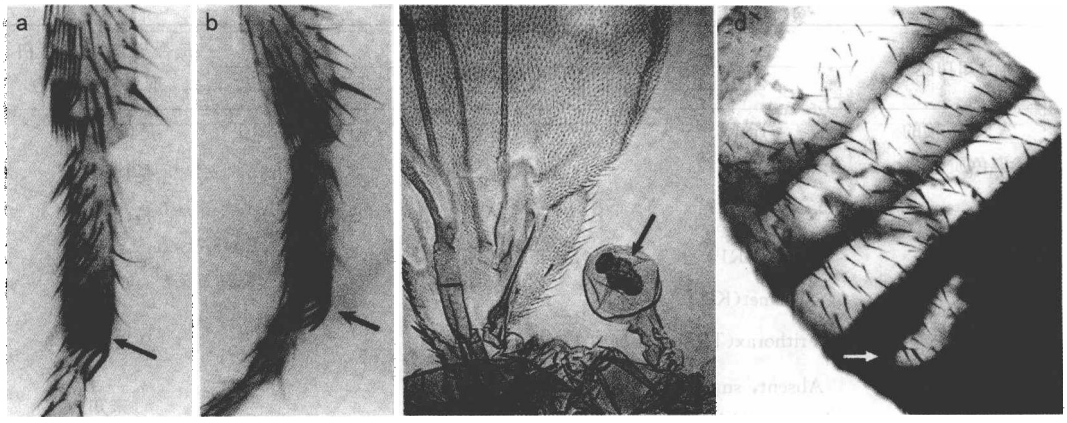



图 12-3　果蝇 trxG 基因突变致发育中细胞命运改变举例

（a）野生型果蝇前足。箭头所示为前足特有的性梳。（b）*Scr* 转录减少导致 *kis* 基因将前足部分转化为中足，虽然转化不完全，但已经表现出性梳齿数量减少（箭头所示）。（c）*Ubx* 转录减少导致 *mor* 基因将平衡器官部分地转化为翅膀（箭头所示）。（d）*AbdB* 表达减少导致第五腹节 *kis* 基因使部分组织（箭头所示）转化，更具有腹部前节特征，具体表现为此体节失去了特征性深色斑点。（a、b、d，经许可转印自 Daubresse et al. 1999）。

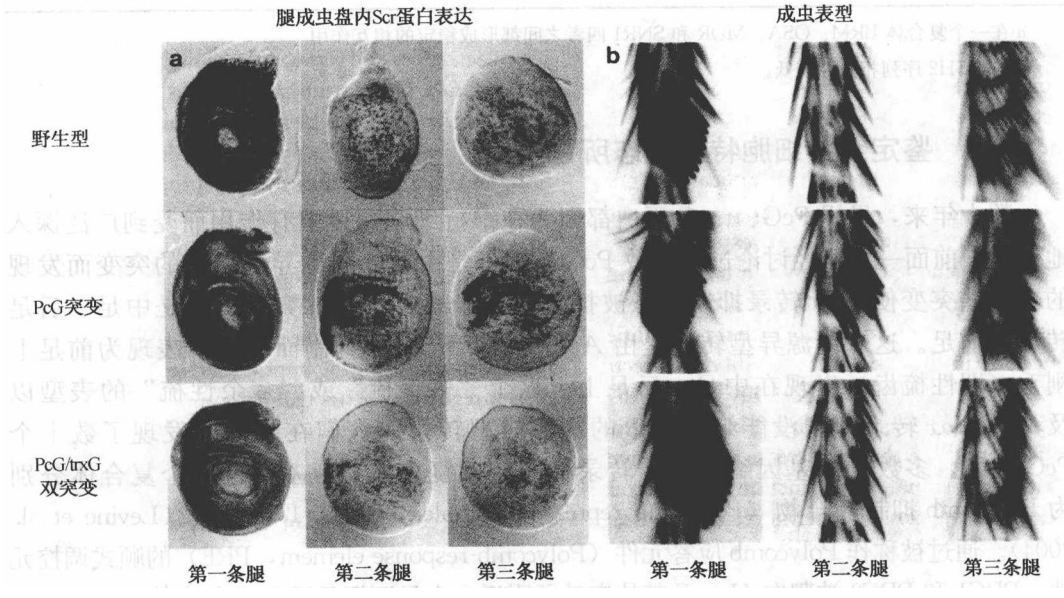



图 12-4　trxG 基因突变阻断了 PcG 突变体中 Hox 基因的去抑制

（a）腿成虫盘经 *Hox* 基因 *Scr* 蛋白抗体标记后，显示出第一胸节（包括前足）和唇的特征。（b）野生型、突变体成虫腿的基跗节。注意野生型成虫只有前足有性梳，而中后足都没有。PcG 基因杂合突变体第二三腿节的 *Scr* 基因去抑制，导致中后足出现了异位的性梳，而正常情况下其中的 *Scr* 基因是沉默的。*brm* 及许多其他 trxG 基因突变可以抑制以上表型（a，经许可转印自 Tamkun et al. 1992 [© Elsevier]；b，经许可转印自 Kennison 2003 [© Elsevier]，有部分改动）。

表 12-1 trxG 蛋白的生化功能

| 已知功能 | 物种 | | | 是否与非 trxG 蛋白形成复合体? |
|---|---|---|---|---|
| | 果蝇 | 人类 | 酵母 | |
| ATP 依赖的染色质重塑 | BRM | BRG1/HBRM | Swi2/Snf2, Sth1 | 是(5～10)[a] |
| | OSA | BAF250 | Swi1、Adr6 | 是(5～10) |
| | MOR | BAF155, BAF170 | Swi3, Rsc8 | 是(5～10) |
| | SNR1 | hSNF5/INI1 | Snf5, Sfh1 | 是(5～10) |
| | Kismet(KIS) | CHD7 | — | 未知 |
| 组蛋白甲基转移酶 | Trithorax(TRX) | MLL1, MLL2, hSET1 | Set1 | 是(5～20) |
| | Absent, small or homeoticl(ASH1) | hASH1 | — | 未知 |
| 调节亚基 | Kohtalo(KTO) | TRAP230 | Srb8 | 是(13～24) |
| | Skuld(SKD) | TRAP240 | Srb9 | 是(13～24) |
| 转录因子 | Trithorax-like(TRL) | BTBD14B | — | 否 |
| 生长因子受体 | Breathless(BTL) | FGFR3 | — | 未知 |
| 其他 | Sallimus(SLS) | Titin | | 未知 |
| | ASH2 | hASH2L[b] | Bre2 | 是(5～20) |

a 在一个复合体 BRM、OSA、MOR 和 SNR1 四者之间都形成稳定的相互作用。

b 与 ASH2 序列相似性较低。

## 1.1 鉴定维持细胞特定状态所需基因

数十年来，果蝇 PcG、trxG 蛋白都因为在维持细胞命运中有作用而受到广泛深入地研究。前面一章已经讨论过，多数 PcG 基因都是由导致同源异型转化的突变而发现的，这些突变使 *Hox* 转录抑制状态被打破。一个经典的 PcG 突变表型是中足、后足转化为前足。这个同源异型转化是由 *ANT-C* 基因 *Scr* 去抑制而形成，表现为前足上刚毛（即性梳齿）出现在中足、后足上。这个“多聚梳”或“多余性梳”的表型以及其他 *Hox* 转录抑制没能维持导致的同源异型转化使人们在果蝇中发现了数十个 PcG 基因。多数 PcG 基因编码两个转录抑制相关复合体的亚基，这两个复合体分别为 Polycomb 抑制复合物（Polycomb repressive complex, PRC）1 和 PRC2(Levine et al. 2004)。通过被称作 Polycomb 应答元件（Polycomb-response element, PRE）的顺式调控元件，PRC1 和 PRC2 被靶向 *Hox* 及其他启动子附近。大量证据显示 PcG 复合体通过调控染色质结构从而抑制转录（Francis and Kingston 2001; Ringrose and Paro 2004）。

trxG 家族成员之间功能性关系以及 trxG 功能维持细胞命运的机理非常复杂。一个蛋白质维持同源异型基因（trxG 蛋白的遗传学定义）恰当的高水平表达，可能有很多种机制，可以不必仅仅作为转录激活因子或进行表观遗传控制的蛋白。除了可以直接激活转录之外，trxG 起作用还可能通过以下方式：提高直接起作用的激活因子的功能；阻断 PcG 抑制子的功能以及创造一个染色质的“开放”状态，这样便促进数量众多的

其他调控复合体发挥作用。此外，还有下面将要讨论的情况，一些 trxG 蛋白在有些激活相关以及另一些抑制相关基因上都能以复杂机制发挥作用。

两个简明的例子就可以解释 trxG 蛋白可能起到的作用有多复杂。人们认为 ATP 依赖的染色质重塑复合体能提高序列特异性 DNA 结合蛋白结合到染色质上的能力，这些复合体之一就含有 trxG 蛋白 BRM 和 MOR。这种 ATP 依赖的染色质重塑复合体是否因此能通过提高激活子和抑制子二者结合上 DNA 的效率而既促进基因激活又促进基因抑制，这个问题尚未解决。其他相关研究提出了一种假说，一些 trxG 复合体可能主要通过阻断 PcG 抑制子复合体作用发挥功能，这种情况下的 PcG 蛋白抑制作用是缺省状态。因此，在后面的例子可以看到，一些 trxG 蛋白维持活性状态反应了一种间接的作用，而不是直接发挥功能。至于它们需要何种机制来维持决定细胞命运的主控基因处于适当的表达水平，trxG 家族进化上以及各蛋白成员功能的保守性能给出一些答案。

## 1.2　其他物种中 trxG 蛋白

在包括人类在内的哺乳动物中能找到几乎所有果蝇 trxG 蛋白的功能对等物（表 12-1）。遗传学以及生物化学的研究显示了果蝇和哺乳动物蛋白在基因表达和发育中有很高的功能保守性。果蝇 trx 在哺乳动物中的同源物——MLL 蛋白是 trxG 蛋白功能保守性的一个好例子。*MLL* 突变会使 *Hox* 基因的激活状态不能维持而导致小鼠中轴骨产生同源异型转化（Yu et al. 1995，1998）。MLL 和 trx 都起着组蛋白赖氨酸甲基转移酶（histone lysine methyltransferase，HKMT）的作用，人类 MLL 蛋白能部分弥补果蝇 *trx* 功能缺失造成的发育缺陷，这是二者功能同源性的直接证据（Muyrers-Chen et al. 2004）。这样，在进化过程中维持确定状态的机制也是高度保守的。

癌症和其他疾病都可能由可遗传的基因表达状态维持不力造成，因此人类许多 PcG、trxG 基因是原癌基因或肿瘤抑制基因这一点也就不足为怪了。例如，11q23 染色体易位与急性淋巴细胞性白血病（acute lymphoblastic leikemia，ALL）、急性髓细胞白血病（acute myeloid leukemia，AML）的相关性使得人 trxG 基因 *MLL* 被发现。其他哺乳动物 trxG 基因也与多种癌症相关（详见第 23 章）。例如，人的 BRG1 是果蝇 *brm* 对等物，该蛋白与视网膜母细胞瘤肿瘤抑制蛋白有物理上的作用，这种作用若被破坏则在特定人类肿瘤细胞系中会出现细胞分裂增加以及恶性转化的后果（Dunaief et al. 1994；Strober et al. 1996）。与 BRG1 在肿瘤抑制中的作用相一致，该基因杂合的突变小鼠容易发生多种肿瘤（Bultman et al. 2000）。人 *INI1* 是果蝇 trxG 基因——*SNF5* 相关基因 1（SNF5-related gene1，*SNR1*）对等物，如果发生突变也预示着这些个体之后会发生癌症，尤其在恶性横纹肌样瘤（儿童易发的一种侵袭性癌症）中大都发现有这种情况（Versteege et al. 1998）。trxG 蛋白与人类疾病的诸多联系也是研究者了解其作用机理的另一个动机来源。

## 1.3　真核转录中 trxG 蛋白具有的多样化功能

trxG 家族中激活因子是一大类调控蛋白，它们有多种多样的功能。这也反应出真

核转录的复杂性，基因特异性转录激活子之间受高度调控的相互作用、通用转录装置中众多成分以及被转录的 DNA 模板都是这一复杂转录过程的组成部分。转录激活是这样进行的，序列特异性激活蛋白结合到特定序列后，募集通用转录装置，RNA 聚合酶 II 结合到启动子上形成前起始复合物，启动子附近的 DNA 螺旋打开，接着 RNA 聚合酶从启动子处有效分离并起始了整个基因的有效延伸。

维持活性转录状态可能涉及这么多激活所需步骤中的任何一个，因为对于任何特定基因，每个步骤都可能成为限速步骤。真核 DNA 包装成染色质更是让 trxG 蛋白可以在另一个层次上起作用。核小体和其他染色体组分都倾向于阻止通用和基因特异转录因子结合 DNA，并阻止 RNA 聚合酶地延伸作用。染色质结构的改变——包括核小体位置方向和结构的改变——几乎能影响转录中的每一步。

转录所需蛋白质都是维持活性状态需要的。实际上，一些 trxG 蛋白起到一些相对通用的作用，而不仅仅专属于维持特定状态。其他 trxG 蛋白的作用则可能在此过程中非常特异，可能直接反作用于 PcG 抑制或通过 DNA 复制和有丝分裂来维持可遗传的基因活跃状态。对于发育生物学家来说，后一类 trxG 蛋白尤其具有吸引力。

## 2. trxG 蛋白与染色质

遗传学研究表明，trxG 基因在转录和发育中有关键作用，这也激发了人们更加努力地去了解 trxG 基因产物的生化功能。有一个假说认为染色质将成为 trxG 蛋白生物学上相关底物，许多研究都以此为理论基础。所有基因都被包装进染色质，这样便形成了一种致密封闭状态，也可以变成开放状态。人们相信这两种状态都是可以遗传的，这些想法成就了一个简单的假说——trxG 蛋白可能调节染色质结构而影响转录调控。进一步对 trxG 基因进行克隆和测序后发现它们的一些产物与 ATP 依赖的染色质重塑相关，有的与其他物种的核小体组蛋白共价修饰有关，这里面包括酵母。因此，尽管酵母没有 Hox 基因和 PcG 抑制因子，这种生物仍然能为了解 trxG 蛋白在真核转录中可能起到的作用提供重要线索。

第一个将 trxG 与染色体相联系的是发现果蝇 trxG 基因 *brm* 与酵母 SWI2/SNF2 基因高度相关（Tamkun et al. 1992）。SWI2/SNF2 是在筛选与接合型转换（*switch* [*swi*]）以及蔗糖非发酵基因（*sucrose-nonfermenting*[*snf*]）相关基因时发现的，随后又发现很多诱导型酵母基因的激活也需要它（Holstege et al. 1998；Sudarsanam et al. 2000）。*swi2*/*snf2* 基因突变体中会有转录缺陷，而核小体组蛋白突变则抑制了这种缺陷表型，这个早期发现最早提出了 SWI2/SNF2 可能通过对抗染色质抑制而激活转录（Kruger et al. 1995）。20 世纪 90 年代初的生化研究肯定了这个假说；SWI2/SNF2 和其他许多在 *swi2*/*snf2* 筛选中发现的蛋白一样，都在一个大的蛋白复合体（SWI/SNF）中作为亚基起作用，这个复合体使用 ATP 水解的能量提高蛋白质与核小体 DNA 的结合能力（Cote et al. 1994；Imbalzano et al. 1994；Kwon et al. 1994）。SWI2/SNF2 是一个起 ATP 酶作用的亚基，相当于这个染色质重塑机器的“发动机”；SWI/SNF 复合体的其他亚基则是调节其与调控蛋白或染色质底物间相互作用的（Phelan et al. 1999）。

trxG 蛋白 TRX 和 ASH1 中都含有 SET 结构域，这样预示着 trxG 和染色质的另一个联系。最初称有与 Su(var)3-9，zeste 基因增强子（Enhancer of zeste，E(z)）和 TRX 蛋白（后面两个分别为 PcG 和 trxG 成员）同源的一串氨基酸序列的结构为 SET 结构域，20 世纪 90 年代末又发现 SET 家族蛋白具有 HKMT 活性。Su(var)3-9 使 H3K9 甲基化，而 E(z) 甲基化 H3K27(Rea et al. 2000；Levine et al. 2004；Ringrose and Paro 2004)。其他章节也讨论过，H3K9 甲基化促进异染色质组装，而 H3K27 甲基化在 PcG 抑制中需要（详情分别见第 5、11 章）。trxG 蛋白的 SET 结构域暗示了组蛋白尾巴甲基化对维持活性状态可能很重要。

有了这些发现，而且人们也越来越意识到染色质重塑、修饰的酶在转录激活中扮演关键角色，这促使生物化学家去寻找含有 trxG 蛋白的蛋白复合体，并且检验在体外这些复合体对染色质结构上的效应。其他实验测试了另一个影响调控的方法是否成立，即 trxG 蛋白能否直接与转录机器作用。如下所述，这些研究揭示了 trxG 蛋白影响调控的方式，一些可以修饰染色质结构，另一些则与转录机器组分直接发生作用。

## 2.1　trxG 蛋白在 ATP 依赖的染色质重塑中起作用

很多生物学过程都涉及染色质重塑复合体的作用，包括转录抑制和激活、染色质组装、染色质结构更高级调控及细胞分化。染色质重塑中研究最深入的 trxG 蛋白是 BRM 及其人类的对等物——BRG1 和 HBRM。和预想的一样，这些蛋白也作为 ATP 酶亚基在与 SWI/SNF 高度相关的复合体中起作用（Kwon et al. 1994；Wang et al. 1996)。SWI/SNF 复合体含有 8～15 个亚基，而且在进化中高度保守（图 12-5)。这些复合体中每一个的 ATP 酶都能作为独立的亚基起作用。因为 ATP 酶结构域与真正解旋酶很相似，因此，这个蛋白家族长期被认为含有“解旋酶”结构域，但实际上 BRM 相关蛋白从来没显示过解旋酶活性，而似乎用其他机制如引起沿 DNA 的易位来影响染色质结构变化（Whitehouse et al. 2003；Saha et al. 2005)。在这次筛选中还发现了另一个 trxG 基因，*moira*(*mor*)，它编码了果蝇这个 ATP 依赖的染色质重塑复合物的另一个关键成员，而 BRM 和 MOR 同源物直接作用形成人类 SWI/SNF 的功能核心（Phelan et al. 1999)。

SWI/SNF 和其他染色质重塑复合体都使用 ATP 水解的能量来改变核小体的结构或位置方向。一般情况下转录因子和其他调控蛋白在组蛋白阻挡下不能靠近 DNA 序列，染色质重塑复合体却能催化染色质结构改变所需的 ATP 依赖过程而使它们接近目标序列（Polach and Widom 1995；Logie and Peterson 1997)。使蛋白接近特异位点有很多模型，包括使组蛋白沿 DNA“滑行”后该位点进入连接区域；使 DNA 从组蛋白八聚体上圈走；更不可思议的是将整个组蛋白八聚体赶到核内的其他区域（图 12-6)。ATP 依赖的染色质重塑还能导致核小体沿一段 DNA 序列的位置变化，核小体间隔改变以及组蛋白交换进出核小体核心——组蛋白八聚体。在发挥每一个功能时，不同重塑复合体都显出不同喜好性。

更高等的真核生物中 SWI/SNF 复合体也大量存在，如哺乳动物细胞核内就有大约 25 000个 SWI/SNF 复合体的拷贝。生物化学分析说明，与其他 ATP 依赖的染色质重塑复合体相比，SWI/SNF 复合体能使蛋白接近极多核小体位点（Fan et al. 2003)。例

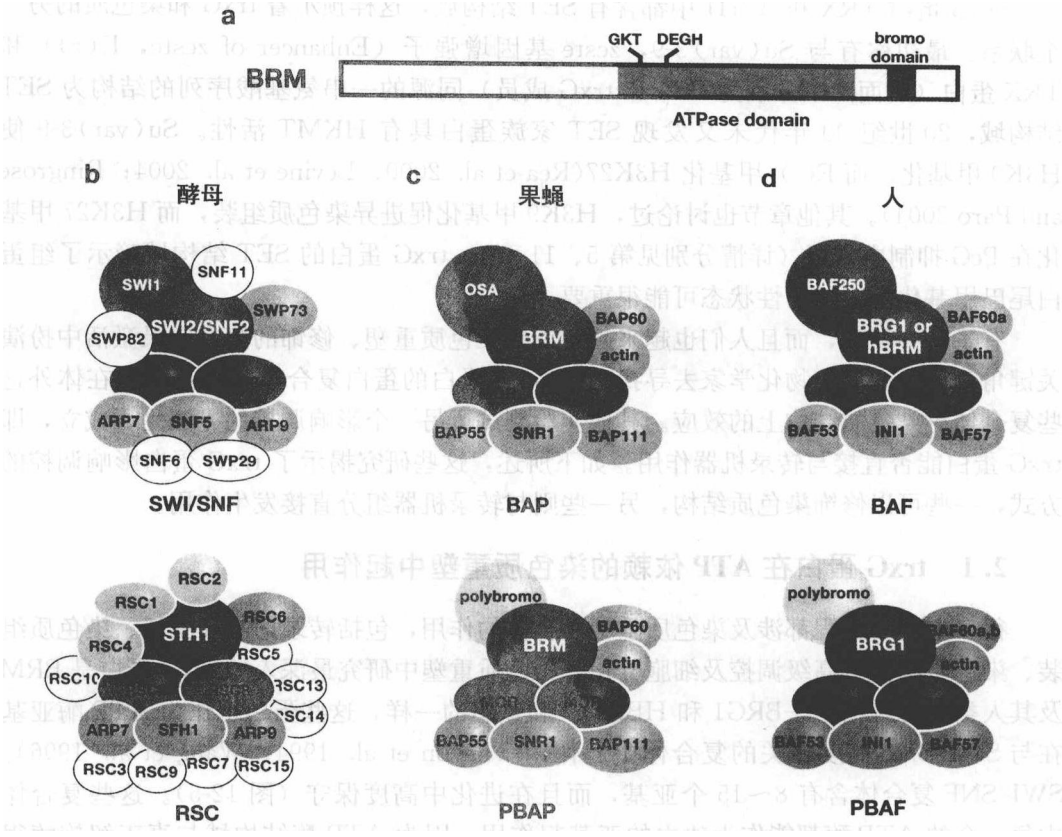



图 12-5 重塑复合体 SWI/SNF 家族

每个复合体都有 SNF2/SWI2ATP 酶家族的一个成员，且含有至少其他 8 个亚基。(a) BRM 蛋白的 ATP 酶结构域、C 端 bromo 结构域（与组蛋白尾巴乙酰化赖氨酸残基有亲和作用）的位置，它们在所有 SNF2/SWI2 家族成员中保守。酵母 (b)、果蝇 (c) 和人 (d) 的 SWI/SNF 复合体。果蝇 trxG 蛋白（BRM、MOR 和 OSA）和其他生物的对等物用彩色标明。这些复合体进一步的信息可以在 Mohrmann 和 Verrijzer 2005 的文献中找到。

如，这之中包括单个核小体中心的位点，此位点两侧有大约各 70bp 不易活动的核小体 DNA，所以要接近它们有很高的能量要求。正在进行的研究（Kassabov et al. 2003）想弄清这是靠与其他重塑复合体不同的超级潜能利用了 ATP 水解能，还是有不同的重塑机制。与其他重塑体不同，SWI/SNF 复合体没有测量出核小体平均间隙的能力；在“交换”H2A/H2B 二聚体上也没有显出相同程度的效力，但在体外它们能做到这一点，且能“赶走”八聚体（Lorch et al. 1999）。其中哪项能力与维持活性状态功能相关至今还不清楚。

SWI/SNF 复合体在所有检验的物种当中都与转录激活相关。这个复合体家族能通过与转录激活子作用靶向基因；能重塑核小体以利于通用转录因子和 RNA 聚合酶Ⅱ最初的结合；并且可以在后期激活过程中被靶向而帮助转录延伸。因此 SWI/SNF 复合体似乎在转录激活的每个步骤中都起作用，而早期能使 RNA 聚合酶Ⅱ结合的步骤应该是

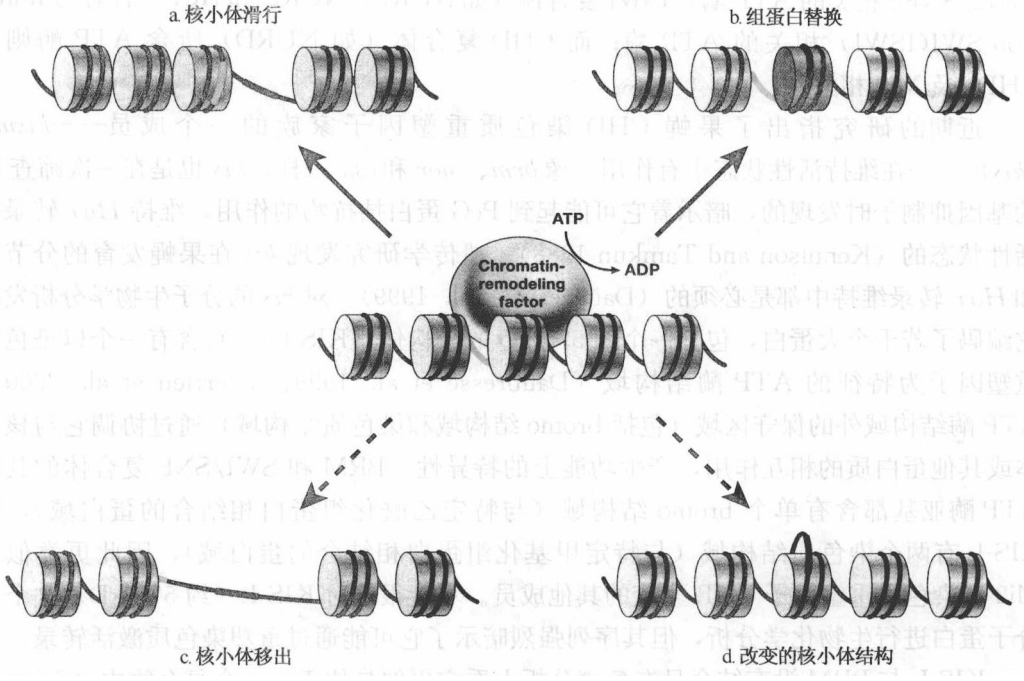



图 12-6　ATP 依赖的染色质重塑机制

通过显示核小体与包裹的 DNA 的相对位置或组成变化说明染色质重塑模型。中间板块是一个初始染色质区域，黄色表示连接 DNA，红色表示核小体 DNA。(a) 核小体在翻译中沿着 DNA 滑行，本来闭合的区域（红色）就暴露出来；(b) 组蛋白变异体取代了正常组蛋白，形成变异核小体；(c) 核小体移出后空出一大段 DNA。这个机制可能依赖于其他蛋白，如组蛋白伴侣或 DNA 结合因子，还有重塑蛋白；(d) 核小体表面空出一个环。SWI/SNF 家族的重塑者可能用了不同机制使核小体中心位点开放，如在核小体表面形成稳定的 DNA 环。

作用重点。酵母中微阵列分析证明 SWI/SNF 复合体不仅能促进激活，还能促使一些基因的抑制（Sudarsanam et al. 2000）。

一个简单的假说可以解释它们在体内如此宽泛的功能，这些重塑体改变核小体结构并因此促进很多种调控因子和复合体的结合及发挥功能。如此，在体外发现的重塑潜质可能反映了它们在体内大量使调控蛋白接近目标的能力。SWI/SNF 复合体可能是仅有能够广泛创造通道的复合体，也正因此其在维持活性状态中很重要。

每个研究的物种都有至少两个不同的 SWI/SNF 复合体，它们都含有 BRM 或高度相关地染色体重塑 ATP 酶。另一种 trxG 蛋白，OSA，是这些复合体间的不同处，一类复合体含有 OSA 而另一类进化保守的复合体含有多聚 bromo 结构域蛋白（图 12-6）（Mohrmann and Verrijzer 2005）。OSA 的生化功能还不明确，但有可能它将所在的 SWI/SNF 复合体靶向到一套特定基因上，这个可能性很具吸引力。

SWI/SNF 并不是真核细胞中仅有的染色质重塑因子，NURF、NURD、ACF 和 CHRAC 等其他几十个不同的染色质重塑复合体也已经被发现（Vignali et al. 2000）。这些复合体可以根据其 ATP 酶特征被分为几个主要组群。SWI/SNF 复合体含有与

SWI2/SNF2 相关的 ATP 酶；ISWI 复合体（如 NURF、NURD 和 ACF）含有与 Imitation-SWI(ISWI) 相关的 ATP 酶；而 CHD 复合体（如 NURD）所含 ATP 酶则与 CHD1 及 Mi2 相关。

近期的研究指出了果蝇 CHD 染色质重塑因子家族的一个成员——*kismet* (*kis*) ——在维持活性状态中有作用。像 *brm*、*mor* 和 *osa* 一样，*kis* 也是在一次筛查 Pc 的基因抑制子时发现的，暗示着它可能起到 PcG 蛋白拮抗物的作用，维持 *Hox* 转录的活性状态的（Kennison and Tamkun 1988）。遗传学研究发现 *kis* 在果蝇发育的分节期和 *Hox* 转录维持中都是必须的（Daubresse et al. 1999）。对 *kis* 的分子生物学分析发现它编码了若干个大蛋白，包括一个约 575 kD 的异构体（KIS-L），它含有一个以染色质重塑因子为特征的 ATP 酶结构域（Daubresse et al. 1999；Therrien et al. 2000）。ATP 酶结构域外的保守区域（包括 bromo 结构域和染色质结构域）通过协调它与核小体或其他蛋白质的相互作用，产生功能上的特异性。BRM 和 SWI/SNF 复合体的其他 ATP 酶亚基都含有单个 bromo 结构域（与特定乙酰化组蛋白相结合的蛋白域），而 KIS-L 有两个染色质结构域（与特定甲基化组蛋白相结合的蛋白域），因此更类似于 Mi2 和染色质重塑因子 CHD 家族的其他成员。尽管很难对 KIS-L（约 575 kD）这个大分子蛋白进行生物化学分析，但其序列强烈暗示了它可能通过重塑染色质激活转录。

KIS-L 与 BRM 没有结合且在色谱分析上看它更像是位于另一个复合物中（Srinivasan et al. 2005）。这两个蛋白却在多线染色体上与彼此又与 RNA 聚合酶Ⅱ高度重合，这又暗示了二者在转录上都起到了相对通用的作用（图 12-7）（Armstrong et al. 2002；Srinivasan et al. 2005）。BRM 功能缺失在转录中阻断一个较早的阶段（Armstrong et al. 2002），而 KIS-L 功能缺失导致处于延伸阶段的 RNA 聚合酶Ⅱ水平减少，而不是起始阶段的水平（Srinivasah et al. 2005）。这些发现暗示了 BRM 和 KIS-L 通过 ATP 依

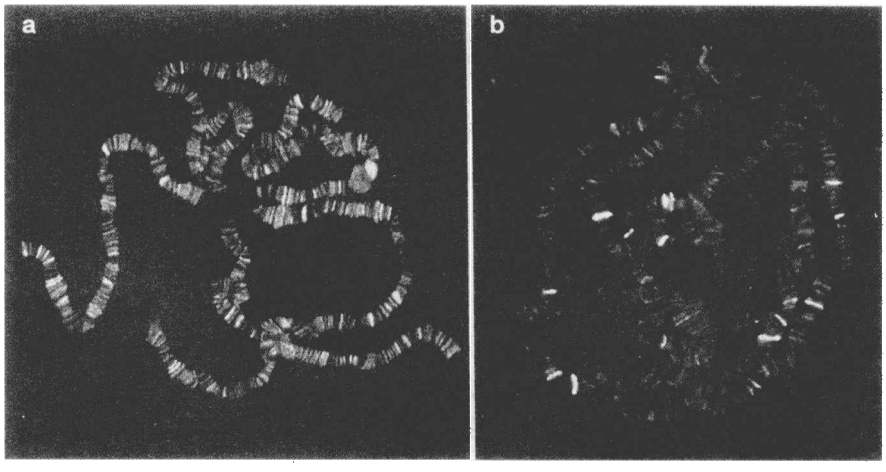



图 12-7 trxG 蛋白在染色质上的分布情况

对果蝇唾腺多线染色体的 BRM（a）或 TRX（b）进行抗体染色可以检验 *trxG* 在基因组范围的分布。BRM 所在的数百个位点与 RNA 聚合酶Ⅱ有极高的重合性，这与在转录激活中相对通用的作用相一致。相反，TRX 只在多线染色体上相对较少的位点有信号。

赖的核小体结构、间距的改变来促进 RNA 聚合酶Ⅱ转录的不同阶段。

未来研究的一个重要问题关系到 ATP 依赖的重塑在维持活性状态中所起的作用。有意思的是四个 trxG 成员有三个（BRM、KIS 和 OSA）已知是 ATP 依赖染色质重塑大复合体的成员，另一个（KIS）也被认为是其成员，但其他众多 ATP 依赖染色质重塑复合体中并没有筛查到果蝇 trxG 蛋白。对此有两个主导的假说，二者并不互斥：一个认为 BRM 和 KIS 染色质重塑复合体靶向发育进程中重要基因；另一个认为它们有维持过程中所需的专门的重塑特性。因此，可能所有活性状态都需要通用的 ATP 依赖重塑功能，而发育中重要基因活性状态的维持碰巧需要这些靶向这些基因的 trxG 成员；也可能是这种维持需要特异的 ATP 依赖功能，这种功能只能由含有 trxG 成员的复合体来完成。

重塑因子用什么机制参与活性状态的表观遗传调控，这一点也很有意义。被考虑到的至少有三类机制。第一类，促进基因特异激活蛋白结合（或复制后的重新结合）有一定间接性，而这种方式需要重塑功能来维持活性转录。这时候，重塑者不是表观遗传机制的“大脑”，而是起着必要的作用，使必需蛋白有效发挥功能。第二类，重塑者或单独或与组蛋白伴侣一起作用，将核小体从一个区域移出，没有核小体占据时，人们认为这个区域会在复制后也维持不被核小体占据的状态。如上面提到的，复制/核小体处理转移装置如何能够精确总调控核小体修饰、核小体位置，这是表观遗传学中一个重要的待解决问题。最后一类，重塑装置能重新调整核小体位置，给出适于激活调控的结构。后一个机制已经有了白蛋白基因研究中得到的证据（Chaya et al. 2001；Cirillo et al. 2002）。肝脏中这个关键基因活性的维持需要几个 DNA 结合因子的参与。其中之一的 FoxA 结合到一个核小体的位点形成了对维持白蛋白基因活性状态至关重要的核小体——FoxA 结构。现在还不清楚肝脏中定位这个特异的核小体是否需要 ATP 依赖的重塑，但这个例子说明特异核小体定位可能起到关键的表观遗传作用。

## 2.2　trxG 蛋白共价修饰核小体组蛋白

调控基因表达的第二个通用工具是对组成核小体的组蛋白 N 端尾巴进行共价修饰。这些从核小体表面伸出来的尾巴，能调节其与其他核小体以及多种结构蛋白、调节蛋白间的作用。组蛋白尾巴上的共价修饰——乙酰化、甲基化、磷酸化等——能协助将调节复合体靶向染色质且能改变尾巴电荷而直接改变核小体压缩成抑制性结构的能力。共价修饰也是维持特定调控状态的一个标记，因为复制后共价修饰的组蛋白能分到两条子链当中，也就把共价标记所含的信息分配到母子细胞当中。复制后组蛋白是否还与一条或两条子链相联系这一点还有争议，大部分因为精确追踪活细胞中个体组蛋白的技术还不成熟，但这一点又对于弄清这个表观遗传调控机制十分关键。有几个 trxG 蛋白能够共价修饰组蛋白尾巴，这些蛋白屡屡被发现在可进行多种修饰的复合体中存在。例如，果蝇 TRX 和在其他生物的对等物将组蛋白 H3K4 甲基化：这个共价标记与很多种生物体内的活性基因息息相关，包括酵母、果蝇和人类。第二个 trxG 蛋白，ASH1（见下文），也有 H3K4 甲基转移酶活性（Beisel et al. 2002；Byrd and Shearn 2003）。测定 H3K4 甲基化在活跃基因上加入和离开的时间过程暗示其在维持酵母活性基因表达的作

用（Santos-Rosa et al. 2002；Pokholok et al. 2005）。发现 trxG 成员具有这种组蛋白修饰活性后，H3K4 标记与活性状态的维持更进一步联系起来。

酵母和人类的 TRX 对等物所在复合体内还有第三种 trxG 蛋白——Ash2，它与 Ash1 没有序列相关性。trithorax 蛋白的酵母同源物——Set1，在一个大约 400 kD 大小的复合体（COMPASS 或称 Set1C）中被发现，这个复合体还有除 Set1 和 Ash2 之外的 5 个蛋白（Miller et al. 2001；Roguev et al. 2001）。这个复合体的唯一已知生化活性是 H3K4 的甲基化，其他蛋白的功能都不清楚，但有的组分可能帮助甲基化标记进行世代传递（见下文）。

人类有三个 TRX 同源物，分别叫做 MLL1、MLL2 和 hSET1，其中 MLL1 蛋白是生化分析中最受关注的。它所在的大复合体（含有超过 10 个成员）有人类 ASH2 的同源物（Hughes et al. 2004；Yokoyama et al. 2004）。这个复合体与酵母复合体一样，也含有一个 WD40 重复蛋白（在人类中叫做 WDR5）（Dou et al. 2005；Wysocka et al. 2005）。最近发现，WDR5 能与 4 号位赖氨酸甲基化的组蛋白 H3 结合（Wysocka et al. 2005），因此，该蛋白绑定到由所在 MLL1 复合体制造的标记上，从而可能加速了该标记的广泛延伸。这与 HP1 的情况很类似，它也能特异性结合到由所在复合体甲基化的 H3K9 标记上（详见第 5、6 章）。

对果蝇和人类研究都为证明含 TRX/MLL 的复合体参与乙酰化提供了依据。人类 MLL 与乙酰基转移酶 MOF 相关，而后者可以乙酰化组蛋白 H4 的 16 号赖氨酸残基，这也是与激活相关的一个常见修饰（Dou et al. 2005）；果蝇 TRX 则与 dCBP 相关，这个乙酰基转移酶起到广泛的特异性激活作用（Petruk et al. 2001）。这样，这些 trxG 复合体发挥的乙酰化和 H3K4 甲基化可能协同作用导致一种激活状态。乙酰化还已知能够阻止导致模板抑制的 H3K9、H3K27 等的甲基化。

另一个 trxG 成员 ASH1 也是个组蛋白甲基转移酶，它能甲基化 H3K4（Beisel et al. 2002；Byrd and Shearn 2003）。目前还不知道任何含 ASH1 的复合体组成，也不了解 ASH1 和含 TRX/MLL1/SET1 的复合体是如何协调发挥作用的，但 ASH1 与乙酰基转移酶 CBP 家族共定位（Bantignies et al. 2000），再一次暗示甲基化、乙酰化有协同作用。

组蛋白共价修饰如何影响 trxG 功能，关于这一点有很多吸引人的未知问题。这些标记起着怎样的作用？共价修饰能通过广泛的机制参与表观遗传调控。甲基化和乙酰化标记可能直接改变染色质致密化过程（有时叫做顺式效应，如图 3-8 所示）。染色质形成致密状态（一般认为这种状态抑制转录）受到组蛋白尾巴上电荷分布的影响，赖氨酸残基上发生的修饰（如乙酰化）能使正常情况下此处的正电荷消除，然后可能直接导致核小体难以形成致密结构，也就提高了模板转录的能力。

有人提出共价标记为转录激活的复合体提供了强结合位点，这些共价修饰在核小体表面形成特异的“结”，此“结”刚好能伸进促进激活的复合体“口袋”，如此增加了这些复合体的结合能和功能发挥。例如，组蛋白尾巴乙酰化提高了与 BRM 同源蛋白的结合，促进乙酰化的模板进行 ATP 依赖的重塑（Hassan et al. 2001）。这种机制常被称作“组蛋白密码”或组蛋白共价修饰的反式效应，可能是表观遗传的中心功能。进一步

研究将要得出 trxG 蛋白制造的哪种标记加强了哪些复合体的结合；要证明绑定到单个修饰残基的能量到哪种程度才能影响功能和靶向；还要知道标记是按什么顺序加上的，还有它们在有丝分裂过程是否能维持。

翻过来看，这个机制则是这些标记能阻止抑制性复合体的结合。抑制性复合体结合必需的关键位点上的共价标记可以强烈抑制此复合体的结合。如抑制性复合体的结合能通过组蛋白 H3 的 K9、K27 甲基化而增加（Khorasanizadeh 2004），这些位点的乙酰化不仅能阻断甲基化的发生，还使组蛋白的“结”变形，不能很好地与抑制性复合体结合。这样来看，共价修饰可以从两个方面影响其他复合体的功能，这个表观遗传调控的潜在形式也就更有潜能。

这些机制不仅不互斥，更可能一起作用协助维持活性状态。化学上增加致密化能力（顺式效应）的标记也能增加复合体的结合（反式效应），进一步促使致密状态形成。反过来，化学上减少致密化的标记也能增加核小体去致密复合体的结合。共价标记改变几种染色质结构特征及调节复合体的绑定能力在机理上运用得如此俭省，这可能为维持活性状态提供强大的工具。

## 3. trxG 蛋白与通用转录装置

trxG 蛋白常常在同一复合体中被发现，skuld(skd) 和 kohtalo(kto) 就是这样两个蛋白。它们是生化上称作 TRAP240(Skuld) 和 TRAP230(Kohtalo) 二者的同源蛋白，这二者都是“协调者”复合体的成员（Janody et al. 2003）。“协调者”复合体很大，作用于基因特异激活因子和含有 RNA 聚合酶Ⅱ的准备起始复合体交界处（Lewis and Reinberg 2003）。因此这些蛋白参与通用的激活过程，大多与 SWI/SNF 重塑者家族方式相同。SKD 和 KTO 可能在维持激活中有一些特殊功能，因为在 trxG 基因筛选实验中协调复合体的其他成分未被识别。SKD 和 KTO 有相互作用，且 *skd kto* 双突变体与任一个单突变体表型相同，这意味着它们可能一起形成了功能性模块，通过一定方式改变了协调者的行为（Janody et al. 2003）。

## 4. 其他 trxG 蛋白的生化功能

其他大多数 trxG 蛋白的生化活性还是个谜。果蝇 trxG 基因 *fsh*（雌性不育同源异型基因，*female-sterile homeotic*）的人类对等物是 Ring3，它编码一个有两个 bromo 结构域的核蛋白激酶，该激酶与细胞周期进程及白血病生成有关，但还不知道它的底物（Denis and Green 1996）。trxG 基因 *Tonalli*(*Tna*) 编码蛋白与 sumo 化 SP-RING 指蛋白相关，暗示它可能也通过共价修饰其他未知蛋白调控转录（Gutierrez et al. 2003）。trxG 基因 sallimus(sls) 是在寻找 Pc 的基因外抑制子时发现的（Kennison and Tamkun 1988），后来又发现它编码了果蝇 Titin 蛋白（Machado and Andrew 2000）。果蝇 Titin 帮助维持肌小节的完整性和弹性，这个作用和其脊椎类对等物一样；另外，染色体凝集和分离中也需要 Titin 这种染色体蛋白（Machado and Andrew 2000）。这些引人注目的

发现暗示了 trxG 蛋白在调节更高级染色质结构时可能的作用。

## 5. trxG 蛋白间功能性相互作用

在了解许多 trxG 和 PcG 成员基本生化活性之后，人们将注意力转移到如何协调它们的活性以调控转录并维持活跃状态上来。虽然缺少研究特定状态的体外系统，但已经有这个问题上的很好进展。一个流行的假说认为，trxG 和 PcG 成员加速了维持活性或抑制状态所需的一串彼此依赖的事件发生。最近对 PcG 复合体 PRC1、PRC2 的研究结

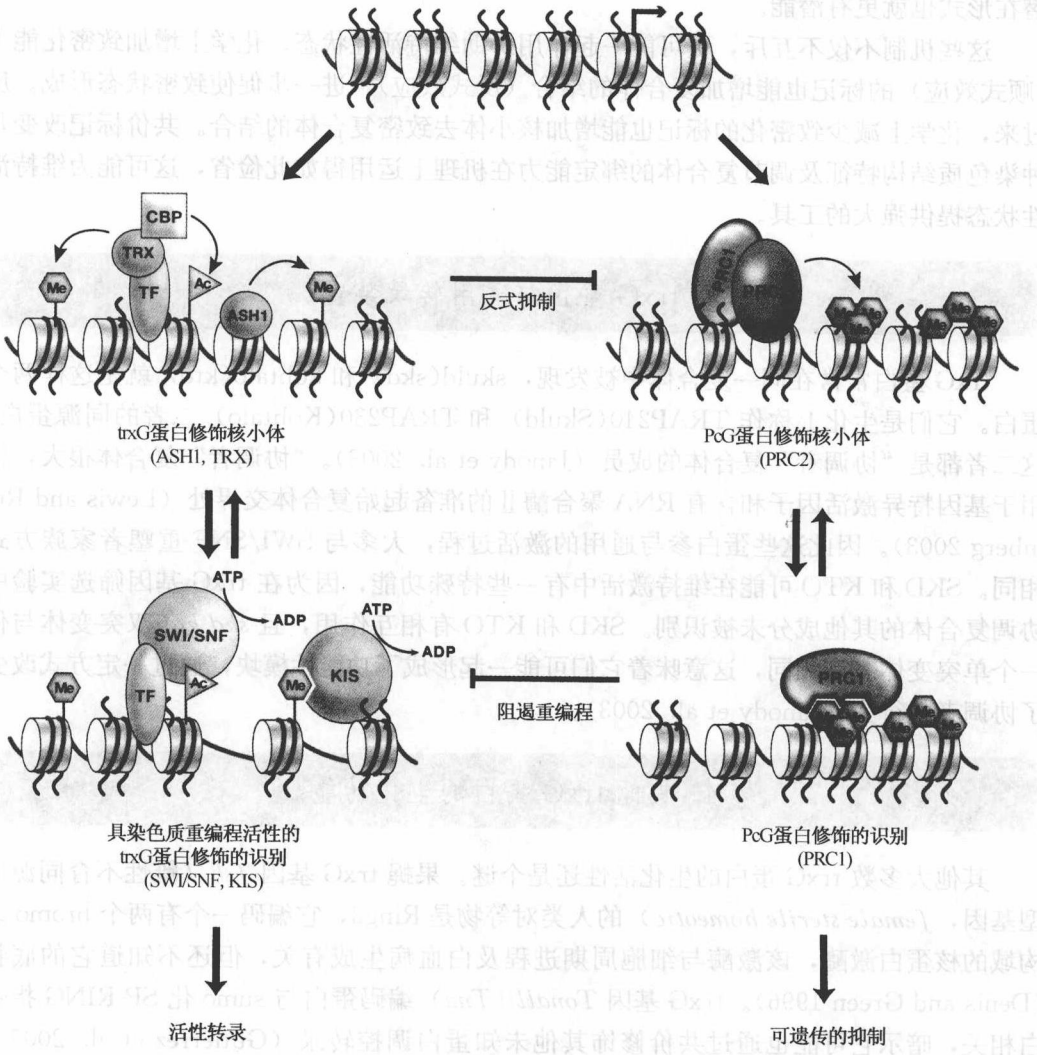



图 12-8 trxG 和 PcG 蛋白的功能及相互作用

trxG 和 PcG 家族都包含可以共价修饰组蛋白和非共价修饰染色质的蛋白质。组蛋白的共价修饰能增加非共价修饰复合体如 SWI/SNF、KIS 或 PRC1 的结合，后者则能导致进一步的共价修饰，这样共价修饰可被识别也可进行循环。

果支持这个想法。PRC2 的组蛋白甲基转移酶亚基 E(z) 将 H3K27 甲基化，产生了一个共价标记，PRC1 的 *Pc* 亚基的染色质结构域就直接能识别此标记（Jacobs and Khorasanizadeh 2002；Min et al. 2003）。因此，一个 PcG 复合体似乎能直接促进另一个 PcG 复合体与染色质的结合。

类似的情况也可能 trxG 家族的组蛋白甲基/乙酰基转移酶（如 TRX 或 ASH1）成员共价修饰核小体，而后直接调控 ATP 依赖的染色质重塑中 trxG 成员（如 BRM［SWI/SNF］或 KIS）的靶向或活性（图 12-8）。与此相一致，BRM 和其他 SWI/SNF 复合体亚基都含有 bromo 结构域，能直接与乙酰化的组蛋白尾巴相作用，同时，KIS 则含有两个染色质结构域（可与甲基化的组蛋白尾巴相作用）。这个模型在最近研究酵母和哺乳类的染色质重塑因子当中有所证明（Agalioti et al. 2000；Hassan et al. 2001），它给出一个机制——可遗传的组蛋白修饰能保持活跃转录所需的组成型的“开放”染色质构象，因此这个模型很有吸引力。

## 6. trxG 蛋白：激活子或是反抑制子

PcG 抑制子和 trxG 激活子之间的功能性关系也是一个重要问题。这些调控蛋白独立发挥其激活或抑制作用，还是它们直接对立作用以维持这种可遗传状态呢？近期有遗传学研究发现，即使没有 TRX 和 ASH1，PcG 复合体移除后基因都将被重新激活（Klymenko and Muller 2004），这就说明有组蛋白甲基转移酶活性的 trxG 蛋白可能以 PcG 阻遏抑制子发挥作用，而不是激活子（图 12-8）。

生化和遗传分析为 trxG 与 PcG 功能间有联系提供了证据。PcG 蛋白具有一个有趣的性质：它与几乎任何 RNA 聚合酶Ⅱ转录的基因靠近时都能抑制转录。trxG 的一些成员（包括 BRM、KIS）在转录中起到广泛作用，另一些则参与染色质重塑，因此很可能是 PcG 抑制子的直接靶点（图 12-8）。主要 PcG 复合体之一的 PRC1 阻断 SWI/SNF 家族重塑复合体的功能很明显是通过阻断复合体靠近模板来实现的（Francis et al. 2001）。有与此一致的观点就是 PcG 抑制的可能机制是阻止 trxG 成员进行 ATP 依赖的染色质重塑。由于直接和 Zeste 蛋白作用，Brahma 复合体和 PRC1 也进一步联系起来，Zeste 蛋白在果蝇基因调控中起着复杂作用，也可能是帮助这两个复合体之间进行直接“对话”的桥梁。

另一个联系 PcG 和 trxG 的蛋白是 GAGA 因子，该因子由 *Trithorax-like* 基因编码，因此也是 trxG 家族的一员（Farkas et al. 1994）。这个蛋白能在一些启动子处起序列特异性激活子作用，也是结合 PRE（见第 8 章）的主要蛋白。PRE 序列引导 PcG 发挥功能，至少一个 PRE 能够在亲和报告基因结构时能作为一个功能性记忆模块，这也更强调出这段序列的重要性。GAGA 因子的结合序列在 PRE 功能发挥中起重要作用，GAGA 蛋白靠近 DNA 可能增强了 PRC1 的结合和功能（Mahmoudi and Verrijzer 2001）。因此 GAGA 因子可能在维持激活和抑制时（分别通过其转录激活性质和与 PcG

蛋白的相互作用）扮演着关键角色。将来的研究还要解决一个问题，就是为什么GAGA和Zeste之类的蛋白看上去既可以与激活性维持装置也可与抑制性装置相作用。

## 7. 结论和展望

关于trxG蛋白功能的两大问题很大程度上还是一些猜测。首先，为什么转录激活所需蛋白中一套相对少的蛋白却在维持活跃状态时有着遗传学上的重要意义呢？是不是因为它们在转录中有通用作用同时表达量却有限？还是在进化中偶然变得对发育关键基因尤其重要了呢？第二，这种活跃状态如何在复制和有丝分裂过程中维持？复制产生两条子链都需要有相同的调控，而有丝分裂需要致密化结构，从而阻止细胞中多数基因的转录。那么是什么样的机制创造的表观遗传标记才能保证有丝分裂后两条子链上基因都会重新激活呢？

大多数trxG蛋白都是在基因表达中作用广泛的复合体的组成部分，这些复合体中很多都含有许多其他非trxG家族的蛋白（表12-1）。这就提出了一个重要问题，是否有一些专门用于维持活性基因表达的功能呢。SWI/SNF重塑因子有特异重塑功能，H3K4甲基化可能指向了特异性复合体和（或）染色质构象，也可能是Skuld/Kohtalo使协调者功能发生了对维持很重要的特异改变。除此之外，还可能每个这样的蛋白都能进行激活各种基因所需的反应，而且这些复合体是在不很吸引的例子中被发现有重要的维持功能（例如，果蝇*Hox*基因表达的相对微小改变也能导致同源异型转化）。现在需要更多信息来了解每个蛋白在激活中采用的精确机制，这样才能解决以上问题。例如，SWI/SNF复合体利用ATP水解的方式是否与其他ATP依赖的重塑复合体相同，还是在它使用此能量改变核小体结构时与后者有某个方面有重要区别？一些结构研究技术，包括结晶学，生物物理技术如单分子分析和荧光共振能量转移（fluorescence resonance energy transfer，FRET）以及体内细节成像技术，都有助于了解维持激活是否有特异表观遗传机制。利用简单模型中trxG复合体上初步进行的功能性研究仅仅是回答这些重要问题的开端。

人们对于维持活跃状态的表观遗传机制现在了解得甚至更少。共价标记的分布是否创造了一种活跃标记？是复制后维持的核小体位置导致染色质“开放”的拉伸还是有特异位置的核小体增加了激活因子的结合？是否trxG发挥功能引起核仁中基因进入转录活跃区域？这些假说都有可能，而且还有更多未想象到的可能性。极复杂的DNA转录装置使得调控和活性转录的表观遗传机制发展有了众多可能采用的方式。这两个领域的交叉有着长久的历史，与转录激活的表观遗传机制一样，很多年来都是一个多产的领域。

朱慧君　李红艳　译
王大亮　校
孙方霖　审校

# 参考文献


Agalioti T., Lomvardas S., Parekh B., Yie J., Maniatis T., and Thanos D. 2000. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-β promoter. *Cell* **103:** 667–678.

Armstrong J.A., Papoulas O., Daubresse G., Sperling A.S., Lis J.T., Scott M.P., and Tamkun J.W. 2002. The *Drosophila* BRM complex facilitates global transcription by RNA polymerase II. *EMBO J.* **21:** 5245–5254.

Bantignies F., Goodman R.H., and Smolik S.M. 2000. Functional interaction between the coactivator *Drosophila* CREB-binding protein and ASH1, a member of the trithorax group of chromatin modifiers. *Mol. Cell. Biol.* **20:** 9317–9330.

Beisel C., Imhof A., Greene J., Kremmer E., and Sauer F. 2002. Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1. *Nature* **419:** 857–862.

Bultman S., Gebuhr T., Yee D., La Mantia C., Nicholson J., Gilliam A., Randazzo F., Metzger D., Chambon P., Crabtree G., and Magnuson T. 2000. A *Brg1* null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol. Cell* **6:** 1287–1295.

Byrd K.N. and Shearn A. 2003. ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. *Proc. Natl. Acad. Sci.* **100:** 11535–11540.

Chaya D., Hayamizu T., Bustin M., and Zaret K.S. 2001. Transcription factor FoxA (HNF3) on a nucleosome at an enhancer complex in liver chromatin. *J. Biol. Chem.* **276:** 44385–44389.

Cirillo L.A., Lin F.R., Cuesta I., Friedman D., Jarnik M., and Zaret K.S. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9:** 279–289.

Cote J., Quinn J., Workman J.L., and Peterson C.L. 1994. Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science* **265:** 53–60.

Daubresse G., Deuring R., Moore L., Papoulas O., Zakrajsek I., Waldrip W.R., Scott M.P., Kennison J.A., and Tamkun J.W. 1999. The *Drosophila kismet* gene is related to chromatin-remodeling factors and is required for both segmentation and segment identity. *Development* **126:** 1175–1187.

Denis G.V. and Green M.R. 1996. A novel, mitogen-activated nuclear kinase is related to a *Drosophila* developmental regulator. *Genes Dev.* **10:** 261–271.

Dou Y., Milne T.A., Tackett A.J., Smith E.R., Fukuda A., Wysocka J., Allis C.D., Chait B.T., Hess J.L., and Roeder R.G. 2005. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121:** 873–885.

Dunaief J.L., Strober B.E., Guha S., Khavari P.A., Alin K., Luban J., Begemann M., Crabtree G.R., and Goff S.P. 1994. The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell* **79:** 119–130.

Duncan I. 1987. The bithorax complex. *Annu. Rev. Genet.* **21:** 285–319.

Fan H.Y., He X., Kingston R.E., and Narlikar G.J. 2003. Distinct strategies to make nucleosomal DNA accessible. *Mol. Cell* **11:** 1311–1322.

Farkas G., Gausz J., Galloni M., Reuter G., Gyurkovics H., and Karch F. 1994. The *Trithorax-like* gene encodes the *Drosophila* GAGA factor. *Nature* **371:** 806–808.

Francis N.J. and Kingston R.E. 2001. Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell. Biol.* **2:** 409–421.

Francis N.J., Saurin A.J., Shao Z., and Kingston R.E. 2001. Reconstitution of a functional core polycomb repressive complex. *Mol. Cell* **8:** 545–556.

Gellon G. and McGinnis W. 1998. Shaping animal body plans in development and evolution by modulation of *Hox* expression patterns. *Bioessays* **20:** 116–125.

Gutierrez L., Zurita M., Kennison J.A., and Vazquez M. 2003. The *Drosophila* trithorax group gene *tonalli* (*tna*) interacts genetically with the Brahma remodeling complex and encodes an SP-RING finger protein. *Development* **130:** 343–354.

Hassan A.H., Neely K.E., and Workman J.L. 2001. Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* **104:** 817–827.

Holstege F.C., Jennings E.G., Wyrick J.J., Lee T.I., Hengartner C.J., Green M.R., Golub T.R., Lander E.S., and Young R.A. 1998. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95:** 717–728.

Hughes C.M., Rozenblatt-Rosen O., Milne T.A., Copeland T.D., Levine S.S., Lee J.C., Hayes D.N., Shanmugam K.S., Bhattacharjee A., Biondi C.A., et al. 2004. Menin associates with a trithorax family histone methyltransferase complex and with the *hoxc8* locus. *Mol. Cell* **13:** 587–597.

Imbalzano A.N., Kwon H., Green M.R., and Kingston R.E. 1994. Facilitated binding of TATA-binding protein to nucleosomal DNA. *Nature* **370:** 481–485.

Jacobs S.A. and Khorasanizadeh S. 2002. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295:** 2080–2083.

Janody F., Martirosyan Z., Benlali A., and Treisman J.E. 2003. Two subunits of the *Drosophila* mediator complex act together to control cell affinity. *Development* **130:** 3691–3701.

Kassabov S.R., Zhang B., Persinger J., and Bartholomew B. 2003. SWI/SNF unwraps, slides, and rewraps the nucleosome. *Mol. Cell* **11:** 391–403.

Kaufman T.C., Seeger M.A., and Olsen G. 1990. Molecular and genetic organization of the antennapedia gene complex of *Drosophila melanogaster*. *Adv. Genet.* **27:** 309–362.

Kennison J.A. 1995. The Polycomb and trithorax group proteins of *Drosophila*: Trans-regulators of homeotic gene function. *Annu. Rev. Genet.* **29:** 289–303.

———. 2003. Introduction to Trx-G and Pc-G genes. *Methods Enzymol.* **377:** 61–70.

Kennison J.A. and Tamkun J.W. 1988. Dosage-dependent modifiers of Polycomb and Antennapedia mutations in *Drosophila*. *Proc. Natl. Acad. Sci.* **85:** 8136–8140.

Khorasanizadeh S. 2004. The nucleosome: From genomic organization to genomic regulation. *Cell* **116:** 259–272.

Klymenko T. and Muller J. 2004. The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5:** 373–377.

Kruger W., Peterson C.L., Sil A., Coburn C., Arents G., Moudrianakis E.N., and Herskowitz I. 1995. Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. *Genes Dev.* **9:** 2770–2779.

Kwon H., Imbalzano A.N., Khavari P.A., Kingston R.E., and Green M.R. 1994. Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex. *Nature* **370:** 477–481.

Levine S.S., King I.F., and Kingston R.E. 2004. Division of labor in Polycomb group repression. *Trends Biochem. Sci.* **29:** 478–485.

Lewis B.A. and Reinberg D. 2003. The mediator coactivator complex: Functional and physical roles in transcriptional regulation. *J. Cell Sci.* **116:** 3667–3675.

Logie C. and Peterson C.L. 1997. Catalytic activity of the yeast SWI/SNF complex on reconstituted nucleosome arrays. *EMBO J.* **16:** 6772–6782.

Lorch Y., Zhang M., and Kornberg R.D. 1999. Histone octamer transfer by a chromatin-remodeling complex. *Cell* **96:** 389–392.

Machado C. and Andrew D.J. 2000. D-Titin: A giant protein with dual roles in chromosomes and muscles. *J. Cell Biol.* **151:** 639–652.

Mahmoudi T. and Verrijzer C.P. 2001. Chromatin silencing and activation by Polycomb and trithorax group proteins. *Oncogene* **20:** 3055–3066.

Miller T., Krogan N.J., Dover J., Erdjument-Bromage H., Tempst P., Johnston M., Greenblatt J.F., and Shilatifard A. 2001. COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci.* **98:** 12902–12907.

Min J., Zhang Y., and Xu R.M. 2003. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* **17:** 1823–1828.

Mohrmann L. and Verrijzer C.P. 2005. Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta* **1681:** 59–73.

Muyrers-Chen I., Rozovskaia T., Lee N., Kersey J.H., Nakamura T., Canaani E., and Paro R. 2004. Expression of leukemic MLL fusion proteins in *Drosophila* affects cell cycle control and chromosome morphology. *Oncogene* **23:** 8639–8648.

Petruk S., Sedkov Y., Smith S., Tillib S., Kraevski V., Nakamura T., Canaani E., Croce C.M., and Mazo A. 2001. Trithorax and dCBP acting in a complex to maintain expression of a homeotic gene. *Science* **294:** 1331–1334.

Phelan M.L., Sif S., Narlikar G.J., and Kingston R.E. 1999. Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. *Mol. Cell* **3:** 247–253.

Pokholok D.K., Harbison C.T., Levine S., Cole M., Hannett N.M., Lee T.I., Bell G.W., Walker K., Rolfe P.A., Herbolsheimer E., et al. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122:** 517–527.

Polach K.J. and Widom J. 1995. Mechanism of protein access to specific DNA sequences in chromatin: A dynamic equilibrium model for gene regulation. *J. Mol. Biol.* **254:** 130–149.

Rea S., Eisenhaber F., O'Carroll D., Strahl B.D., Sun Z.W., Schmid M., Opravil S., Mechtler K., Ponting C.P., Allis C.D., and Jenuwein T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406:** 593–599.

Ringrose L. and Paro R. 2004. Epigenetic regulation of cellular memory by the polycomb and trithorax group proteins. *Annu. Rev. Genet.* **38:** 413–443.

Roguev A., Schaft D., Shevchenko A., Pijnappel W.W., Wilm M., Aasland R., and Stewart A.F. 2001. The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J.* **20:** 7137–7148.

Saha A., Wittmeyer J., and Cairns B.R. 2005. Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. *Nat. Struct. Mol. Biol.* **12:** 747–755.

Santos-Rosa H., Schneider R., Bannister A.J., Sherriff J., Bernstein B.E., Emre N.C., Schreiber S.L., Mellor J., and Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419:** 407–411.

Simon J. 1995. Locking in stable states of gene expression: Transcriptional control during *Drosophila* development. *Curr. Opin. Cell Biol.* **7:** 376–385.

Simon J.A. and Tamkun J.W. 2002. Programming off and on states in chromatin: Mechanisms of Polycomb and trithorax group complexes. *Curr. Opin. Genet. Dev.* **12:** 210–218.

Srinivasan S., Armstrong J.A., Deuring R., Dahlsveen I.K., McNeill H., and Tamkun J.W. 2005. The *Drosophila* trithorax group protein Kismet facilitates an early step in transcriptional elongation by RNA Polymerase II. *Development* **132:** 1623–1635.

Strober B.E., Dunaief J.L., Guha S, and Goff S.P. 1996. Functional interactions between the hBRM/hBRG1 transcriptional activators and the pRB family of proteins. *Mol. Cell. Biol.* **16:** 1576–1583.

Sudarsanam P., Iyer V.R., Brown P.O., and Winston F. 2000. Whole-genome expression analysis of snf/swi mutants of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **97:** 3364–3369.

Tamkun J.W., Deuring R., Scott M.P., Kissinger M., Pattatucci A.M., Kaufman T.C., and Kennison J.A. 1992. *brahma*: A regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68:** 561–572.

Therrien M., Morrison D.K., Wong A.M., and Rubin G.M. 2000. A genetic screen for modifiers of a *kinase suppressor of* Ras-dependent rough eye phenotype in *Drosophila*. *Genetics* **156:** 1231–1242.

Versteege I., Sevenet N., Lange J., Rousseau-Merck M.F., Ambros P., Handgretinger R., Aurias A., and Delattre O. 1998. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394:** 203–206.

Vignali M., Hassan A.H., Neely K.E., and Workman J.L. 2000. ATP-dependent chromatin-remodeling complexes. *Mol. Cell. Biol.* **20:** 1899–1910.

Wang W., Cote J., Xue Y., Zhou S., Khavari P.A., Biggar S.R., Muchardt C., Kalpana G.V., Goff S.P., Yaniv M., et al. 1996. Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15:** 5370–5382.

Whitehouse I., Stockdale C., Flaus A., Szczelkun M.D., and Owen-Hughes T. 2003. Evidence for DNA translocation by the ISWI chromatin-remodeling enzyme. *Mol. Cell. Biol.* **23:** 1935–1945.

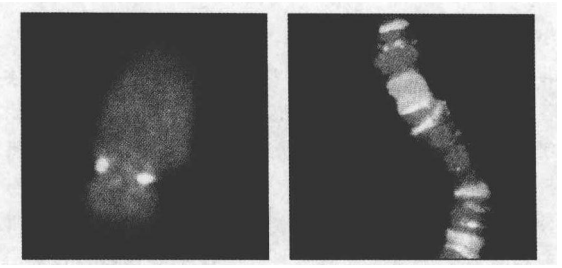
Wysocka J., Swigut T., Milne T.A., Dou Y., Zhang X., Burlingame A.L., Roeder R.G., Brivanlou A.H., and Allis C.D. 2005. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121:** 859–872.

Yokoyama A., Wang Z., Wysocka J., Sanyal M., Aufiero D.J., Kitabayashi I., Herr W., and Cleary M.L. 2004. Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol. Cell. Biol.* **24:** 5639–5649.

Yu B.D., Hanson R.D., Hess J.L., Horning S.E., and Korsmeyer S.J. 1998. MLL, a mammalian *trithorax*-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proc. Natl. Acad. Sci.* **95:** 10632–10636.

Yu B.D., Hess J.L., Horning S.E., Brown G.A., and Korsmeyer S.J. 1995. Altered Hox expression and segmental identity in *Mll*-mutant mice. *Nature* **378:** 505–508.

# 第 13 章



# 组蛋白变体和表观遗传学

Steven Henikoff[1] and M. Mitchell Smith[2]

[1] *Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109-1024*

[2] *Department of Microbiology, University of Virginia, Charlottesville, Virginia 22908*

## 本章目录

1. 所有生物的 DNA 都被结构性蛋白所包装
2. 真核生物核心组蛋白由古细菌组蛋白进化而来
3. 大多数组蛋白在 DNA 复制后被置入
4. 组蛋白变体的置入发生于整个细胞周期
5. 着丝粒被一个特定的 H3 变体标记
6. 置换性组蛋白变体 H3.3 存在于活跃的染色质区域
7. H2AX 的磷酸化在 DNA 双链断裂修复中具有功能
8. H2AZ 作用于转录调控
9. H2A 变体组装和置换的蛋白复合物
10. 其他功能未知的使染色质具有差异的 H2A 变体
11. 许多组蛋白进化以更紧密地包装 DNA
12. 结束语和未来研究

参考文献

## 概要

组蛋白通过装入核小体核心部分以包装DNA，使DNA双螺旋结构环绕其上。在进化中，组蛋白折叠结构域蛋白已从其古细菌祖先分化为我们所熟知的真核生物核小体中八聚体的四种亚基。组蛋白进一步多样化为各种变体，产生了具有表观遗传学效果的染色质分化。对组蛋白变体进化、结构和代谢的研究，为理解染色质在细胞进程和表观遗传学记忆中的作提供了基础。

绝大多数组蛋白在S期合成以迅速置入复制叉经过后留下的原有组蛋白间的空隙。另外，组蛋白变体对常规的S期组蛋白的不依赖于复制的置换，可以潜在地使染色质得以区分。

因某种组蛋白变体引起的染色质分化特别显著地体现于着丝粒，该处特殊的核小体中装有H3变体CENP-A，形成了动粒组装的基础（见本章题图左侧）。所有真核生物中都含有对应于CENP-A的着丝粒H3，CenH3。在植物和动物中，着丝粒处含CenH3核小体的可靠组装似乎并不需要着丝粒DNA序列，这是一个令人印象深刻的表观遗传例子。一些CenH3在接触DNA的区域已经适应性地进化，这暗示了着丝粒间的相互竞争，而CenH3和其他着丝粒特异DNA结合蛋白则相应改变以适应。这一进程也许可以解释植物和动物中巨大而复杂的着丝粒。

着丝粒以外，染色质也可以被一类组成型表达的H3(H3.3）的置入而区分。H3.3是复制非依赖性核小体组装的底物。H3.3的置换发生于活跃基因（题图右侧，绿色显示的是一个果蝇染色体上的H3.3)，是一个具有潜在表观遗传学意义的动态进程。H3和H3.3之间的区别加上共价修饰可能承载了活跃转录基因座处染色质属性的变化。

几种H2A变体也可以区分或调控染色质。H2A. X含有4个氨基酸的的一个特定羧基端模块的变体，其丝氨酸残基是DNA双链断裂处的磷酸化位点。H2AX磷酸化是DNA双链断裂修复中的一个早期事件，被认为可以招募修复所需装置的组分。H2AX磷酸化也在哺乳动物精子发生过程中标记失活的XY联会染色体，并且为浓缩、配对和可育性所必需。

H2AZ这个在结构上有很大差异的变体长久以来是一个谜。酵母中的研究暗示了H2AZ在建立转录能力和拮抗异染色质沉默中的作用。将核小体中H2A置换为H2AZ的生化复合物是一个ATP依赖性的核小体重塑因子，是该多样化的染色质结合装置家族中成员具有一个特定功能的首例。

两个脊椎动物特异性变体，macroH2A和$H2A^{Bbd}$，在体外装入核小体时显示了相反的特性，前者妨碍转录而后者促进转录。这些特性和它们在表观遗传失活的哺乳动物X染色体上的定位模式相一致，Xi富含macroH2A而缺少$H2A^{Bbd}$。

这些研究逐渐形成这样的观点：组蛋白变体和将其装入核小体的进程提供了一种可能作为表观遗传进程基础的染色质基本分化形式。

## 1. 所有生物的 DNA 都被结构性蛋白所包装

与细胞器 DNA 相比，核 DNA 巨大的长度显然需要紧密包装，因此演化出了结构性蛋白以达到这一目的。第一级缩短了双螺旋并保护其免于损伤，但在每个细胞周期允许 DNA 聚合酶充分接触每个碱基对。另外，这些结构性蛋白有利于更高级别的折叠以进一步降低染色体长度。很可能是因为包装 DNA 所需的严谨条件，在所有的细胞生命形式中只发现两种类型的结构性蛋白（Malik and Henikoff 2003）。细菌 DNA 被 HU 蛋白包装，真核生物 DNA 则由组蛋白，而古细菌 DNA 由 HU 蛋白或组蛋白。

组蛋白将 DNA 包装成核小体颗粒，这个结构使组蛋白占了真核生物染色体质量的一半。尽管如此，组蛋白还在基因表达、染色体分离、DNA 修复和真核生物中其他基本的染色体有关进程中起着多种作用。这些染色体有关的进程的特殊要求，导致了不同的组蛋白变体的进化。一种组蛋白变体装入一个核小体代表了染色质潜在的深刻变化。确实，最近的工作显示一些组蛋白变体是由不同的核小体组装复合物放置的，这暗示了染色质至少在部分上因为组蛋白变体的装入和置换而多样化。

四种核心组蛋白具有变体的倾向不同。例如，人只有一种 H4 亚型但是有多种具有不同属性和功能的 H2A。显然，核心组蛋白在核小体颗粒中不同的位置使其面对不同的进化压力，引起重要的 H2A 和 H3 多样化，而非 H2B 和 H4。通过众多可提供基因组序列的真核生物，我们可以得出结论：这些多样化发生在真核生物进化的不同时期。尽管如此，祖先组蛋白折叠蛋白多样化为四种核心组蛋白的过程发生在真核生物核进化的早期或者早于核的进化。通过考虑这些古老的事件，我们可以深入了解使其继续多样化为目前状态变体的原因。

## 2. 真核生物核心组蛋白由古细菌组蛋白进化而来

真核生物核小体是一个复杂的结构，含有缠绕了约两圈 DNA 的由四种核心组蛋白组成的八聚体，同时组蛋白末端和连接组蛋白介导了核心部件外的多种相互作用（Arents et al. 1991；Wolffe 1992；Luger et al. 1997）。古细菌的核小体简单得多，显然它们类似于真核生物核小体的原始结构（Malik and Henikoff 2003）。古细菌核小体中 DNA 只缠绕一圈于不含末端尾巴的组蛋白折叠结构域蛋白组成的四聚体上。通过比较结构可以看出古细菌和真核生物核小体的亲缘关系：古细菌四聚体的骨架几乎和 $(H3 \cdot H4)_2$ 四聚体一样（图 13-1）。当以古细菌核小体体外组装成染色质时，得到的纤维与以 $(H3 \cdot H4)_2$ 四聚体组装而来的相类似。因此，普遍认为古细菌核小体在四聚体两端各添加上 H2A · H2B 二聚体并加上组蛋白末端尾巴后，演化成为真核生物核小体，以使得 DNA 能缠绕第二圈。另外，DNA 围绕古细菌组蛋白核心缠绕成右手超螺旋，而在真核生物中是左手超螺旋。

对真核生物核小体起源的进一步深刻了解来自对古细菌核小体亚基结构的研究。尽管多数古细菌组蛋白是未分化的单体或者组成四聚体的结构上可互换的变体，但有些则

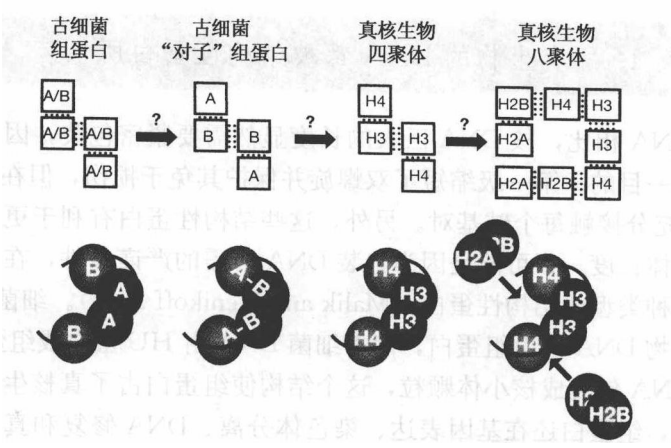



图 13-1 古细菌组蛋白祖先进化为真核生物核小体的模型

含有可相互交换的亚基 A 和 B 的古细菌四聚体可能是二聚体或融合状态的二聚体（“对子”）。接下来的变化可能是一个基因分开产生真核生物四聚体中的 H3 和 H4，形成只裹有一圈 DNA 的 $(H3 \cdot H4)_2$ 四聚体。H2A 和 H2B 可能来自相似的事件，并组装在四聚体上下两侧以容纳 DNA 的第二圈缠绕（图中未显示）。图中单个的点表示二聚体间的结合，而双点表示相邻二聚体间含四个螺旋的束状结构（重印自 Malik and Henikoff 2003）。

形成头尾相接的融合二聚体并进一步形成这种融合二聚体的二聚体（图 13-1）。当两个这样的融合二聚体装入核小体时，二者处于结构上有区别的地方。占据着核小体中不同的位置，古细菌融合二聚体的成员分别演化，使其适应于核小体中单一的位置。相反，可互换位置的单体则无法对特定位置进行适应。的确，古细菌二聚体的两个成员在双方不同源的地方发生多样化。这提供了一种可能的场景来解释祖先组蛋白折叠结构域蛋白如何分化成真核生物核小体中占据不同位置的四种不同的亚基。和推测出来的古细菌祖先一样，真核生物组蛋白形成二聚体，其中 H2A 和 H2B 形成二聚体而 H3 则和 H4 形成二聚体（并在溶液中以四聚体形式稳定存在）。因此，尽管四种真核组蛋白间和与古细菌组蛋白之间缺乏显著的序列相似性，二聚体结构的惊人重叠性暗示了真核组蛋白从更简单的古细菌组蛋白进化和分化而来。

似乎起源于古细菌前后串联二聚体的 H2A · H2B 和 H3 · H4 二聚体的不对称性，可能引导了接下来向真核组蛋白变体的多样化。H2A 和 H3 与古细菌串联组蛋白二聚体的第一个成员一致，二者在真核生物进化中发生了多次多样化。相反，类似第二个成员的 H2B 和 H4 罕有功能上的多样性。八聚体中 H3 和 H2A 都存在同源二聚相互作用（图 13-2），而 H4 和 H2B 则只和其他组蛋白接触。故而，H2A 或 H3 发生同源二聚作用处残基的改变可以抵制混合八聚体的形成，也允许含有 H2A 或 H3 变体的核小体独立于母本核小体而进化。例如，H3 接触面的四螺旋束状结构 DNA 缠绕在核小体上发生左手超螺旋（Arents et al. 1991；Luger et al. 1997），而古细菌核小体上的 DNA 超螺旋是右手性的（Marc et al. 2002）。显然，一个 H3 祖先中四螺旋束状结构的突变是形成这种逆转的原因。通常来说，促进亚基独立进化的结构特性可能是核小体结构多样化的先决条件。

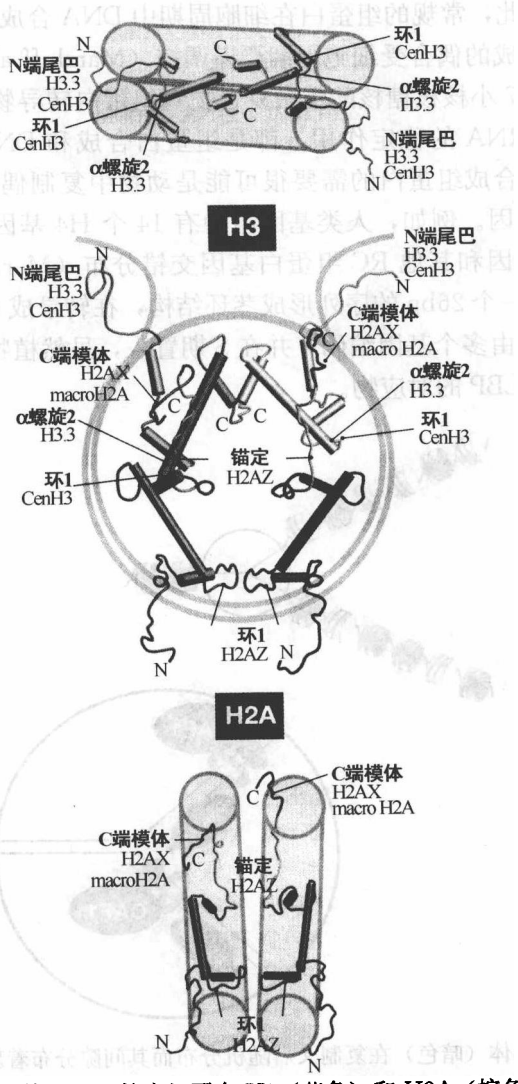



图 13-2　核小体核心组件中组蛋白 H3（蓝色）和 H2A（棕色）的位置

H3 变体间差异的四个残基经黄色显示（经许可重印自 Henikoff and Ahmad 2005）。

虽然我们可以理性地推出从古细菌串联二聚体到真核生物核心组蛋白的演化，其他基本问题依然存在。组蛋白尾巴从何而来？H2A・H2B 是何时加入的？这些事件发生在真核生物细胞核进化之前、之时还是之后？为什么所有已知的古细菌核小体含有缠着一圈 DNA 的四聚体，而真核生物核小体含有缠着两圈 DNA 的八聚体？为什么超螺旋手性发生变换？很可能更多古细菌或者来自早期真核生物的序列可以显示出中间形式以回答这些问题。

## 3. 大多数组蛋白在 DNA 复制后被置入

为了将真核细胞中几乎所有的 DNA 组装入核小体，需要在 DNA 复制时将染色质

拷贝（图 13-3）。因此，常规的组蛋白在细胞周期中 DNA 合成期（S 期）产生。S 期组蛋白合成与 DNA 合成的偶合受细胞周期严谨调控（Marzluff and Duronio 2002）。这在动物中尤其显著：U7 小核核糖核酸蛋白复合物对组蛋白转寻物的特殊加工和茎环结合蛋白（SLBP）对 mRNA 的稳定作用，都是组蛋白合成和 DNA 复制严密协调的一部分。S 期快速和大量合成组蛋白的需要很可能是动物中复制偶合（RC）的多种组蛋白基因呈簇状分布的原因。例如，人类基因组中有 14 个 H4 基因，多数分布在两个簇状结构中，其中 H4 基因和其他 RC 组蛋白基因交错分布（Marzluff et al. 2002）。动物中，RC 组蛋白 3′端一个 26bp 的序列形成茎环结构，在转录成 mRNA 时被 SLBP 识别。常规的植物组蛋白也由多个基因编码，并在 S 期置入，虽然植物组蛋白转寻物被多腺苷化并且似乎不含有 SLBP 的对应物。

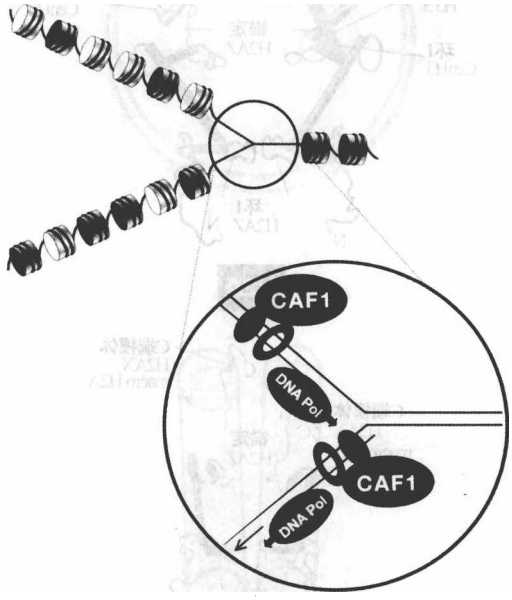



图 13-3 旧核小体（暗色）在复制叉后随机分布而其间隙分布着新核小体（淡色）。图中放大显示的是 CAF1 介导的前导链和跟随链处的核小体组装。DNA 聚合酶（绿色）；复制持续性夹 PCNA（蓝色环）；组蛋白 H3·H4 四聚体（粉红）；新合成的 DNA（红色）。

由于表观遗传可能来自于一种染色质“状态”的传递，复制偶合核小体装配的进程一直是关注的焦点。这一进程的生化过程随着体外实验中在复制中 DNA 上组装核小体体系的发展得以阐述。这些研究显示一个三亚基复合物，组蛋白组装因子 1（CAF-1），作为一个组蛋白伴侣蛋白促进核小体组装第一步 H3·H4 的置入（Loyola and Almouzni 2004）。CAF-1 和复制持续性夹 PCNA 相作用，意味着 DNA 复制和复制偶合的组装在近距离发生。酿酒酵母中的工作显示，体外 RC 组装中复合物的亚基无一是生长必需的，暗示体内存在 RC 组装的冗余机制。许多酵母染色质以复制非依赖性（RI）方式组装的事实（Altheim and Schultz 1999）为这种显著的冗余性提供了一种解释。下文将讲到，组蛋白变体通常通过复制非依赖性核小体组装而被装入。

RC 组装在酿酒酵母中并非完全冗余。一个有趣的发现是，大 CAF-1 亚基的缺失导致端粒处表观遗传沉默的丢失（Loyola and Almouzni 2004）。RC 组装和表观遗传沉默的关系已经扩展到了拟南芥上，CAF-1 亚基的缺失导致多种可归因于表观遗传记忆缺失的缺陷。虽然这些现象的机理基础尚不清楚，似乎明确的是，复制叉后新核小体的装入对维持一种表观遗传沉默的状态很重要。

核小体表观遗传信息继承的一个先决条件是复制后子染色单体上必须分布有原有的核小体（图 13-3）。确实如此，广泛的研究已经显示原有的核小体是完整而明显随机地传到子染色单体（图 13-3）（Annunziato 2005）。尽管如此，该一继承性进程还很不清楚，同样不清楚的还有新组蛋白是如何获得表观遗传信息的。一个流行的模型是，新核小体因其邻近于旧核小体而被修饰（Jenuwein 2001）；但是，这一假想的进程尚缺乏证据支持，而其他延续某种表观遗传状态的方式必须被考虑（Henikoff and Ahmad 2005）。表观遗传信息如何遗传到子代细胞仍然是生物中一个大而未决的问题。对组蛋白变体及其置入机理的研究也许能提供一些线索。

## 4. 组蛋白变体的置入发生于整个细胞周期

正如我们已经看到的，核心组蛋白可以根据它们的祖先序列和在核小体中位置加以划分。连接体组蛋白具有特征性的带翼螺旋结构域，而非组蛋白折叠结构域，且该结构域结合于核小体间的连接 DNA(Wolffe 1992)。虽然存在常规组蛋白的少量变体，它们和主要形式的组蛋白似乎是可交换的。例如，哺乳动物 H3. 1 和 H3. 2 仅差一个氨基酸，而且不清楚是否给这两种变体带来不同的生物学属性。为在 S 期置入大量常规组蛋白，组蛋白多基因的存在是真核生物基因组的典型特征。普遍存在和占绝对多数的常规 S 期组蛋白导致了对组蛋白变体的忽视，直到最近才得到改善。

对组蛋白变体兴趣的复兴部分来自于认识到它们不同于常规 S 期组蛋白，可以带来染色质的显著分化。一点不同在于它们置入染色质的方式。复制耦合组装在基因组范围内把新核小体装入旧核小体间的空缺位置，而复制非依赖性组装涉及所在处核小体或其亚基的置换（Marzluff et al. 2002）。因此，复制非依赖性组装可以通过用变体替换常规组蛋白，以切换染色质状态。将一个组蛋白加以替换也可以同时擦除或改变翻译后修饰的模式。因而，复制非依赖性组装有可能重置由组蛋白及其修饰介导的表观遗传学状态。近来对组蛋白变体及其置入进程的研究引发了对表观遗传信息承继和重塑的新理解。以下，我们讨论特定组蛋白变体的一些方面，这些方面都对染色质分化有作用而且可能涉及表观遗传信息的传递。

## 5. 着丝粒被一个特定的 H3 变体标记

着丝粒是真核生物染色体的定义性特征，是有丝分裂纺锤体微管结合位点。酿酒酵母（*Saccharomyces cerevisiae*）着丝粒最早在分子细节上得以描述，其上 125bp 是着丝粒形成的必要和充分条件（Amor et al. 2004a）。尽管如此，植物和动物的着丝粒很不

一样，一般含有上兆个碱基的短重复序列。不同于酿酒酵母，这些复杂的着丝粒 DNA 序列的作用并不确定，因为功能完全的人的新着丝粒在完全缺乏类似着丝粒重复序列的异位区域也能自动形成（图 13-4）。这些和其他证据反驳了 DNA 在决定着丝粒位置时的直接作用（见第 6 章）。

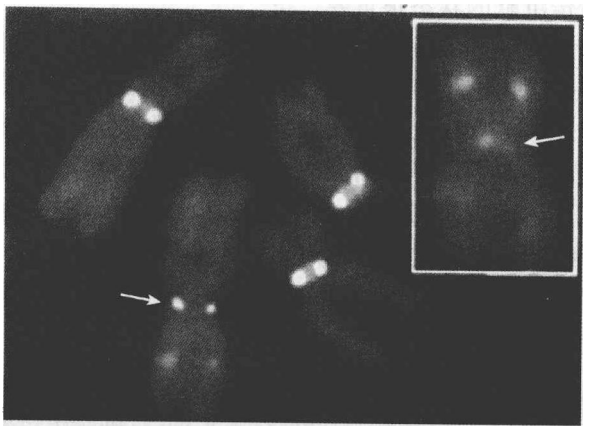


图 13-4 人新着丝粒（箭头所指）缺少着丝粒的 α 卫星 DNA，但是含有 CENP-A 和异染色质绿色的抗 CENP-A 染色和红色的抗 CENP-B（标记 α 卫星 DNA）确定了一个第四号染色体的新生着丝粒缺少 α 卫星 DNA（主图）。该染色体在其他方面是正常的，已经在正常个体内至少有丝分裂三代了。小图显示抗 HP1 染色，表明在缺少卫星 DNA 的情况下，异染色质依然形成于新着丝粒周围（经许可重印自 Amor et al. 2004b [©国家科学院]）。

对着丝粒特征和继承性的一个重要见解来自组蛋白 H3 变体 CENP-A（题图）的鉴定。CENP-A 特异性位于着丝粒并且替换了核小体中 H3 的位置（Palmer et al. 1991）。值得一提的是，在精子发生时组蛋白转换为精蛋白的过程中，尽管几乎所有其他组蛋白都丢失了，CENP-A 依然和着丝粒相结合（Palmer et al. 1990）。CENP-A 研究中的这个早期发现暗示 CENP-A 可能是雄性基因组中着丝粒特性的一部分。这一见解的普遍性并没有得到充分肯定，直到发现相对于 DNA 序列，CENP-A 是一个更好的着丝粒标记（Amor et al. 2004a），而且 CENP-A 的对应物在所有真核生物基因组中都能找到（图 13-5）（Malik and Henikoff 2003）。因此，虽然酿酒酵母的着丝粒由 125bp 的共有序列决定，该位置也是含有 Cse4 着丝粒（CenH3）变体的着丝粒核小体所在处。裂殖酵母（*Schizosaccharomyces pombe*）中，CenH3 核小体阵列占据着两侧均为呈异染色质特性的 H3 核小体的着丝粒中央核心区域（Amor et al. 2004a）。在果蝇和脊椎动物中，CenH3 和 H3 阵列交替分布，展示一种独特的组蛋白修饰模式（Sullivan and Karpen 2004）。这种差异可能是因为着丝粒只占据了中期染色体着丝粒压缩区域的外侧边缘（题图）。这和线虫中观察到的“全着丝粒”染色体一致，其中微管广泛接触后期染色体，而 CenH3 则沿着染色体分布在引导侧（图 13-5 右侧）（Malik and Henikoff 2003）。的确，在所有检查过的真核生物中，独特的 CenH3 变体精确的标记了着丝粒。这种显著的普遍性和所有真核生物中完成有丝分裂时着丝粒的存在，提出了这样一个可能性：最初的常规 H3 起源于 CenH3。

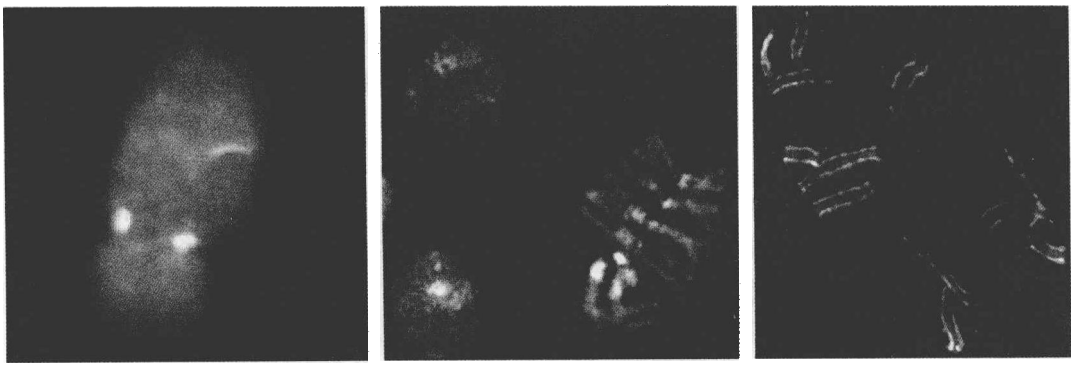


图 13-5　模式真核生物中的着丝粒 H3 变体

（左）人染色体经抗着丝粒特异性组蛋白 H3 变体 CENP-A 染色（绿色）和标记 α 卫星 DNA 的抗 CENP-B 染色（红色）（Peter Warburton 许可图片）。（中）黑腹果蝇抗 CenH3 抗体（红色）染出了中期染色体以及整个间期的着丝粒（Suso Platero 许可图片）。（右）线虫抗 CenH3 抗体（绿色）染出了前期染色体从一端一直延续到另一端的全着丝粒（红色）（Landon Moore 许可图片）。

多种真核生物中的遗传学实验证明了 CenH3 在动粒形成和染色体分离中的重要性（Amor et al. 2004a）。因为 CenH3 在整个细胞周期中都留在原位，它们形成了有丝分裂和减数分裂中其他动粒蛋白组装的基础（见第 6 章）。染色体研究中一个未解决的问题是这些蛋白如何连接着丝粒和纺锤体微管，形成可以支撑后期染色体分开时作用于动粒的强大拉力的结构。在酵母中已发现有几十种动粒特有蛋白（详见第 6 章），虽然并不清楚它们如何与含 CenH3 核小体以及 CENP-C 等其他结构蛋白作用。另一挑战来自如何描述将 CenH3 装入核小体的进程。着丝粒只是染色质总量一小部分的事实妨碍了用生化手段研究这个突出的问题，但是我们期望技术进步会带来对动粒结构和动态的更好理解。

CenH3 的进化不同于任何其他的组蛋白类型。尽管组蛋白 H3 的序列基本不变，这反映自然选择对每个残基超乎寻常的纯化能力，但是 CenH3 进化很快，特别是在植物和动物品系中（Malik and Henikoff 2003）。从氨基端尾巴看来最为明显，不同分类学群组的 CenH3 长度和序列相差如此之大以至于无法放在一起进行比对。甚至 CenH3 的组蛋白折叠结构域的进化速度也比 H3 的高几个数量级。是什么原因导致作用于着丝粒的 H3 和作用于其他地方的 H3 之间有如此惊人的进化差异？

果蝇和拟南芥 CenH3 基因快速进化的区域显示出远高于预期的核苷酸替换频率（Malik and Henikoff 2003）。这是适应性进化的典型特征。在植物和动物中另一个主要的着丝粒结构蛋白，CENP-C，亦可见适应性进化（Talbert et al. 2004）。虽然在有遗传冲突的基因中，如宿主和寄生者之间的军备竞赛，适应性进化已被广为报道，但 CenH3 和 CENP-C 是所有物种中已知的具适应性进化的单拷贝基因中仅有的必需基因。在 CenH3 和 CENP-C 中，适应性进化的区域对应于 DNA 结合和定位区域。这暗示主要着丝粒结合蛋白向进化中的着丝粒 DNA 适应，以允许着丝粒区域染色体结合保守的动粒装置，从而连接到纺锤体微管。有人提出在雌性减数分裂时，着丝粒相互竞争以进

入卵子而非作为极体被丢失（Talbert et al. 2004）。为了实现着丝粒的扩张，会发生军备竞赛，可能是通过不对称姐妹染色单体交换。通过 CenH3 和 CENP-C 对这个减数分裂驱动过程的主体性抑制导致在 DNA 作用区域更多的替换变化。在酿酒酵母等没有机会进行着丝粒竞争的生物中，不存在这种着丝粒扩张倾向，这可能是它们着丝粒小而且 CenH3 和 CENP-C 蛋白经过了很强选择的原因。

所以，我们看到着丝粒这一基因组中特殊的区域由一个组蛋白变体家族得以区分，其序列是一场导致着丝粒如此复杂的军备竞赛的遗迹。每个细胞周期中将新的 CenH3 核小体组装到着丝粒的复制非依赖性（RI）组装过程仍然未知（Amor et al. 2004a）。着丝粒核小体显著地缺少序列特异性，但是它们不仅可以忠实地定位到完全不同于着丝粒的新着丝粒（图 13-4），而且酵母同源物 Cse4 可以有效地功能性替换人的 CENP-A（Wieland et al. 2004）（这两者均无适应性进化；Talbert et al. 2004）。数千万年来，我们的着丝粒异乎寻常的留在同一位置，却没有任何明显的维持这一过程的决定序列。考虑到表观遗传学指的是不依赖于 DNA 序列的继承性，着丝粒位置的长久继承是可想象的最极端形式。然而，我们仍然在寻找一种机理以解释它们如何能在哪怕一个细胞周期中得以维持（第 14 章会更深入讨论）。

## 6. 置换性组蛋白变体 H3.3 存在于活跃的染色质区域

和着丝粒一样，转录活跃的染色质被认为通过表观遗传得以维持，而活跃染色质富含一种 H3 变体，H3.3（Henikoff and Ahmad 2005）。H3.3 序列上非常类似常规的 H3，只相差四个氨基酸。如此少的区别，可能会让人以为这两种形式是可以互换的。尽管如此，在果蝇中，H3.3 通过复制耦合或复制非依赖性核小体组装完成置入，而 H3 只通过复制耦合的形式在复制叉处置入。这两个变体间的差别由其蛋白本身所编码，H3 和 H3.3 之间四个区别中的三个显然和阻止 H3 以复制非依赖性方式组装有关（α 螺旋 2 中，图 13-2）。可溶性的人组装复合物的纯化确定了这两种形式介入不同的组装进程：H3.1 在复制耦合组装中和 CAF-1 共纯化，而 H3.3 和其他成分包括 HirA 共纯化出来，参与 RI 组装。

虽然四个氨基酸的差异看起来显得很不显著，但考虑到人，果蝇和蛤具有相同的 H3.3 序列时，这些与 H3 的差异显得非常突出。种系发生分析显示 H3/H3.3 在植物、动物/真菌、鞭毛虫和孢子虫的真核进化中，至少有四次独立的进化（Malik and Henikoff 2003）。尽管有着不同于动物和真菌的来源，动物中的 H3/H3.3 对和植物中（称为 H3.1[RC] 和 H3.2[RI] ——为避免混淆，我们称所有的 RC 亚型为 H3，而所有的 RI 亚型为 H3.3）的惊人的相似。在果蝇中妨碍 RI 置入的同一氨基酸簇也在植物中发现并存在差异，同时动物中剩余的区别点（H3 第 31 位是 Ala，而 H3.3 是 Ser 或 Thr）也存在于植物中。真菌特别的有趣。祖上，它们含有 H3 和 H3.3；然而，子囊菌，包括酵母和霉菌（mold），丢失了 H3。因此，动物中最引人注目的 H3 专一性 RC 形式甚至不存在于酵母中。

对整个染色质中 H3.3 的研究显示其富集于活跃区域（Henikoff and Ahmad

2005)。尽管如此，在组蛋白修饰可以把活跃和沉默染色质区分开这一现象极大的兴奋了染色质领域的时期，各种因素使 H3.3 这种潜在的活跃染色质“标记”显得含糊和难于理解。首先，没有已有抗体可以有效区分染色质中的 H3 和 H3.3(87～90 位在核小体中被 DNA 螺旋覆盖)，而针对许多不同翻译后修饰的优秀抗体都已经有了。另外，H3 和 H3.3 间看上去细微的差异并未暗示其能给染色质带来质的区别，而组蛋白修饰则主要发生在已知影响染色质相互作用或结合染色质相关蛋白的末端赖氨酸上。四膜虫中，S 期 H3 可以替换其对应的置换形式，这确认了这两种 H3 亚型可以互换的认识。最后，影响广泛的“组蛋白密码”假说把核小体想象为染色质分化时修饰酶的固定靶标(Jenuwein and Allis 2001)。但是，越来越明显的是，染色质是高度动态的，即便异染色质结合蛋白的结合时间也只有一分钟或更短暂 (Phair et al. 2004)。似乎活跃转录基因的染色质处于持续的动态，伴有连续的组蛋白置换 (Henikoff and Ahmad 2005)。H3 和 H3.3 之间的三个主要区别使 H3.3/H4 二聚体成为 RI 组装的底物，而 RI 组装本身对染色质影响深远。这一进程的结果是，活跃转录的区域被 H3.3 标记 (图 13-6)。该进程证据来自在 RNA 聚合酶 I 和 II 转录的基因座，第 9 位赖氨酸甲基化的 H3 被复制非依赖性 (RI) 置换成加标签的 H3.3(Schwartz and Ahmad 2005)。

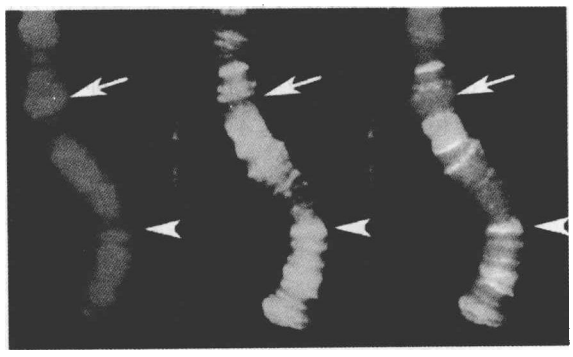


图 13-6　H3.3 倾向于分布在果蝇多线染色体的活跃转录区域

(左) DAPI 染色 (红色) 显示 DNA 带型模式。(中) H3.3-GFP (绿色) 位于间带 (interband)，指示 RNA 聚合酶 II 的位置。(右) 叠加以后的效果 (经许可重印自 Schwartz and Ahmad 2005)。

活跃基因座处染色质的动态天性导致原有组蛋白修饰的去失。这提供了一个可能的方法，解释 H3K9 和 H3K27 高度甲基化 (通常和沉默染色质相关的组蛋白修饰) 的沉默染色质如何能被激活。研究显示组蛋白上的甲基和组蛋白本身一样稳定 (Waterborg 1993)，虽然最近发现的单和双甲基化 H3K4 特异性去甲基化酶 (Shi et al. 2004) 说明组蛋白上某些甲基可以通过酶促反应被移除。大体上，组蛋白共价修饰模式可能来自组蛋白被组装时它们上面的既有修饰。这样说来，修饰酶会跟随着组装装置，并可能进而促进这一进程 (Henikoff and Ahmad 2005)。

这一动态组装模型预测富集于活跃染色质的组蛋白修饰也应该富集于 H3.3，在植物和动物中对总 H3 和 H3.3 上修饰的测量数据正是如此。更进一步，从这一模型出发，预期 H3 和 H4 乙酰化以及 H3K4 和 H3K79 的甲基化相互之间会有很强的相关性，

正如在多种系统中所观察到的（O'Neill et al. 2003；Kurdistani et al. 2004；Schubeler et al. 2004）。最后，活跃基因处动态的复制非依赖性（RI）组装可以解释为何 CAF-1 突变引起沉默作用的减弱（Loyola and Almouzni 2004）：只有约 10%的酵母基因组被认为处于沉默状态，而这可能是酵母基因组中仅有的不处在动态替换中的部分。缺少 CAF-1 介导的复制耦合（RC）组装的情况下，RI 组装会发生在整个酵母基因组，激活原来沉默的区域。可能在多细胞动物中出现某种专一于 RC 组装的 H3 变体是一种适应的结果：以在细胞中保持大部分染色质处于表观遗传沉默状态。

被不同修饰的 H3.3・H4 二聚体置换暗示了一个在分裂细胞中继承活跃染色质的简单模型（Henikoff and Ahmad 2005）。通过 RC 组装装入的常规核小体，与原有 RI 组装核小体混杂后，活跃染色质在这种稀释后仍保持活跃，只要这些混杂的核小体没有阻碍转录起始和延伸等活跃进程。转录活性的继续会在下一细胞周期恢复染色质状态，维持发育过程中活跃染色质的永久延续。对于 CenH3，该组蛋白变体由一个 RI 组装进程维持的可能性也存在，因为它能插入在细胞分裂后期染色体分离时的张力下形成的普通核小体间的空隙中。

当细胞脱离细胞周期开始分化，它们不再产生或装入 S 期组蛋白，同时 H3.3 开始累积。例如，400 天大鼠的大脑中，H3.3 占到了 H3 总量的 87%（Henikoff and Ahmad 2005）。染色质这种逐渐的替换有无功能意义尚不清楚。我们同样不知道这个允许替换发生的主动过程是否和转录活跃基因座处发生的一样。一种可能性是，经过的 RNA 聚合酶或者染色质重塑装置对染色质的打开引起了所在处核小体的解开和H3.3・H4 二聚体的偶然丢失（图 13-7）。紧随着聚合酶会发生核小体的重组装，通过 HirA 复合物用 H3.3・H4 二聚体替换丢失的二聚体。只有当聚合酶分布太紧密而不能发生组装时，核小体才会被完全解开。

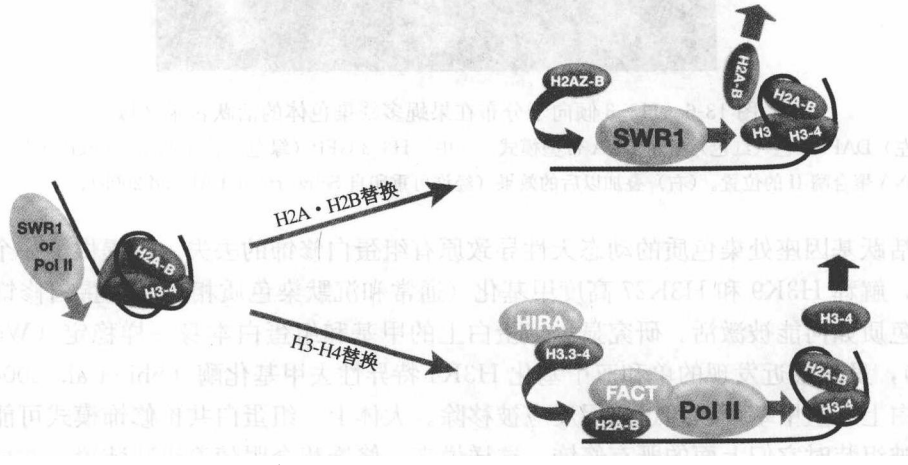



图 13-7 复制非依赖性置换或交换的模型

一个大的分子机器（SWR1 复合物或 RNA 聚合酶Ⅱ）在行进中部分或完全解开核小体。结果可能是异源二聚体亚基被保留，例如 FACT 介导的将 H2A・H2B 从 RNA 聚合酶之前转移到后方（Formosa et al. 2002；Belotserkovskaya et al. 2003），也可能是异源二聚体丢失。在后一种情况中，染色质修复由 H2AZ・H2B（上）或 H3.3・H4（下）补充丢失的二聚体。

## 7. H2AX 的磷酸化在 DNA 双链断裂修复中具有功能

所有真核生物组蛋白 H2A 也是由不同变体组成的一个家族。变体 H2AX 的特征是羧基端氨基酸序列，SQ(E 或 D)Θ，Θ 代表一个疏水氨基酸。该序列模块中的丝氨酸是磷酸化位点，产生一个修饰后被称为“γ-H2AX”的蛋白。染色质的动态属性，和 H2AX 磷酸化，在 DNA 双链断裂时尤为明显（Morrison and Shen 2005）。即使单个双链断裂也可能致死，这使得该损伤需要马上被修复以恢复双螺旋的连续性。双链断裂的检测发生在其形成的一分钟左右，这随即触发断裂位置紧连处的 H2AX 快速磷酸化。该磷酸化由类磷酸肌醇 3 激酶激酶（PI3K-like kinase）家族成员完成。在这一起始事件之后，H2AX 的磷酸化沿着染色体快速蔓延，标记了断裂处附近相对大的一个染色质区域。最后，双链断裂最终通过同源重组或非同源性末端连接完成修复，然后磷酸化标记被去除。

H2AX 磷酸化并非双链断裂的检测或修复所必需的，因为该基因的删除或者目标丝氨酸残基的突变并不终止修复。但是，由于这些突变体修复效率降低，且对放射性伤害和基因毒性物质超敏感，H2AX 并不仅仅是一个损伤的标记。现在认为 H2AX 在双链损伤修复中至少以两种方式起作用。首先，它可能帮助招募或保持在断裂处修复所需的蛋白（Morrison and Shen 2005）。其次，它可能通过募集黏结蛋白，一个保持姐妹染色单体在一起的蛋白复合物，稳定断裂点旁边的染色质（Lowndes and Toh 2005）。

H2AX 的进化与其他组蛋白变体不同。虽然几乎所有真核生物中都有一个 H2AX 基因，但是它具有多个相对近的起源（Malik and Henikoff 2003）。例如，果蝇（Drosophila）中发现的 H2AX 版本不同于另一种双翅类昆虫疟蚊（Anopheles）中的。有可能，从常规的 H2A 进化出一种新 H2AX 仅仅是那个 SQ（E 或 D）Θ 模块带来的结果。在常规 H2A 的羧基端进化出这样一个模块大概在进化历程中重复发生过。因新版 H2AX 的产生导致原有 H2AX 的丢失可能是因为 H2AX 需要均一化分布所造成的，因为双链断裂可以发生在基因组的任何位置。如果一个已有 H2AX 基因发生突变，使得它和常规 H2A 差异过大而导致其组装效率降低或不均一，则会有很强的选择压力将其替换为更类似于常规 H2A 的版本。该原理有助于解释果蝇 H2AX 的例外情况。和其他动物不一样，它不是来自 H2A，而是来自挺远的 H2AZ 世系（描述见下）。如果要成为 H2AX 仅仅是在核小体中占据 H2A 的位置并且具有可磷酸化的羧基末端模块，H2AZ 也可以具有这种能力。

双链断裂修复显然是 H2AX 磷酸化的普遍功能，而且似乎这个进程没有涉及稳定的表观遗传学因素。但是，不含 H2AX 的小鼠是不育的，同时哺乳动物精子发生的细胞学检测显示一个惊人的表观遗传特征，就是 XY 二价体上的 H2AX 被特异性磷酸化（图 13-8）（Fernandez-Capetillo et al. 2003 ）。该染色体对位于减数分裂前期一个叫做“性体（sex body)”的部位上，其在雄性减数分裂过程中可能涉及性相关基因的沉默。H2AX 磷酸化是正常性体形成所必需的，且缺乏 H2AX 的精母细胞无法配对或浓缩，进而无法在减数分裂中失活 X 和 Y 的基因。XY 二价体的磷酸化不同于双链断裂处的

进程。性体处的 XY 磷酸化不需要断裂的存在，而最显著地发生在染色体的未配对区域。H2AX 为何靶向地发生在未配对染色体的机理，以及这一事件如何导致浓缩，配对和沉默尚不清楚。尽管如此，推测该作用可能与其招募及结合黏结蛋白的能力有关是一有趣的想法。

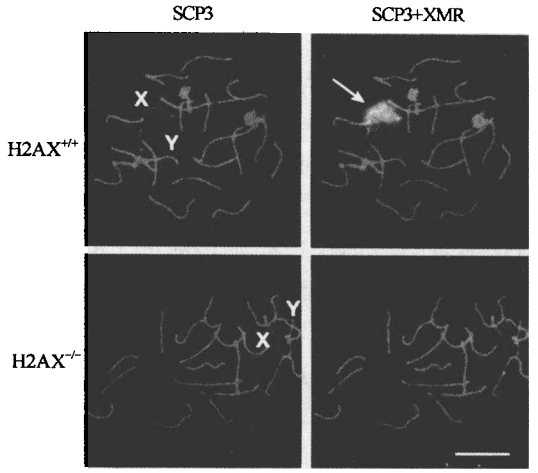



图 13-8 精子发生粗线期显示性体形成对 H2AX 的依赖性

正常哺乳动物精母细胞中，一种称为性体（sex body)的核结构（箭头，右侧绿色标示）围绕着未配对的 XY 二价体（标示于左侧）。细胞接合丝复合物顺着配对的染色体分布（红色）。H2AX 通常富含于性体（$H2AX^{+/+}$）。在 $H2AX^{-/-}$ 的精母细胞中，性体不会形成且一个性体抗原决定簇会分散到各常染色体中（右下图）。标尺为 10μm。图片经 Shantha Mahadevaiah and Paul Burgoyne 允许使用（Fernandez-Capetillo et al. 2003）。

## 8. H2AZ 作用于转录调控

对组蛋白变体兴趣的复兴特别体现在 H2AZ（或 H2A. Z）（Kamakaka and Biggins 2005）。H2AZ 几乎是普遍存在的，它在真核生物进化的早期分化自一种祖先 H2A。与这种独立的世系相一致，酿酒酵母和果蝇中的遗传实验显示组蛋白 H2A 和 H2AZ 都进化以完成独立的非重叠的功能。H2AZ 对多数生物是一种必需组蛋白，从鞭毛原生动物到哺乳动物。尽管如此，芽殖和裂殖酵母中，H2AZ 的删除产生可存活的细胞，虽然完全突变体有多种表型。这些表型促进了在酵母中对 H2AZ 的遗传学和生化特性的描述。

目前测试过的多数生物中，H2AZ 大约占总 H2A 的 10%。它广泛但并非均一的分布在染色体上。这在果蝇多线染色体上可以很好的看到，体现在所呈现的一种特殊带型。使用酵母和小鼠细胞的染色质免疫共沉淀实验结果与这种带型结果一致。虽然 H2AZ 倾向分布于酵母基因的启动子区域，这种特异性并不适用于所有的 H2AZ 存在位点。果蝇中，H2AZ 定位和基因表达之间没有清楚的关系。因此，尽管 H2AZ 置入的机理是已知的（讨论如下），但目前并不了解决定其集中在何处的规则是什么。

多种观察指出 H2AZ 在调控基因表达上有重要作用（Kamakaka and Biggins 2005）。酿酒酵母的突变分析显示 H2AZ 的功能与两类不同的基础转录因子部分重叠，

分别是核小体重塑复合物 Swi/Snf 和组蛋白修饰复合物 SAGA。尽管单独的 H2AZ，Swi/Snf 或 SAGA 功能的缺失都不致死，两条通路的任何组合的同时缺失都是致死的。其他遗传学和生化实验暗示这些角色包括转录起始和延伸中的功能（详见第 10 章）。而且，H2AZ 置入的平衡因其抗沉默因子的角色而联系到了表观遗传学。H2AZ 基因的删除导致沉默染色质从端粒往内部延伸，而这一缺陷可以通过其他沉默因子基因的删除而得到抑制（图 13-7）（见第 4 章）。删除 H2AZ 基因对总体基因表达的影响已经用酵母基因芯片分析过。虽然 H2AZ 缺失突变体中大部分调控基因表达量降低，有相当一些基因的表达量增加。由于还不清楚哪些变化反映了直接调控而哪些是间接的，H2AZ 核小体可能同时正向和负向调控基因表达。进一步，我们也不知道 H2AZ 在转录和异染色质中的多种角色是来自一种统一的机理还是更复杂的多种通路的组合。

与酿酒酵母中的现有图景相反，哺乳动物细胞中 H2AZ 倾向于分布在异染色质区域。确实，H2AZ 和异染色质结合蛋白（Heterochromatin-associated Protein 1，HP1）间有物理作用（Fan et al. 2004）。尽管这可能暗示 H2AZ 在多细胞动物中沉默的作用，但是值得注意的是果蝇异染色质中表达的那部分基因是需要 HP1 参加的（Weiler and Wakimoto 1995）。如果哺乳动物细胞中 H2AZ 的定位反映一个相似的进程，那么这个变体在酵母中已经清楚确立的促进转录和拮抗沉默的角色很可能反映了这个变体普遍的基础属性。

H2AZ 在染色体分离的表观遗传学中有着其他的角色。最开始认识到的表型之一是 H2AZ 缺失突变体在裂殖酵母中有减数分裂染色体分离缺陷。更多近期的实验加强了这种联系。哺乳动物细胞中用 RNA 干涉（RNAi）减少 H2AZ 的实验引起了着丝粒周围 HP1 结合、基因组不稳定和染色体分离错误等缺陷（Kamakaka and Biggins 2005）。相似的，酿酒酵母中 H2AZ 缺失突变体显示减数分裂染色体丢失的增加，且其与编码着丝粒和减数分裂纺锤体组分的基因间有显著的遗传相互作用（Krogan et al. 2004）。仍然完全有可能，H2AZ 对染色体分离的作用是它对基因转录调控的结果。但是，一个引人注意的假设是；染色体分离的机理已经进化到不仅利用 H3 变体，还利用 H2A 变体。

H2AZ 是如何影响转录、沉默、异染色质甚至于染色体分离的呢？H2AZ 核小体的高分辨率结构显示了该变体一些独特的属性（Suto et al. 2000）。与 H2A 核小体相比，H2AZ 在核小体表面呈现更大的酸性结构域，而这个差别可能具有功能意义。例如，“接合结构域”（图 13-2）的一部分与组蛋白 H4 作用，确定了其在果蝇中发挥功能所必需的片段。进一步，突变研究和体外结合实验的结果认为这个多出来的酸性区域对核小体与 HP1 的结合作主要的贡献（Dryhurst et al. 2004）。有趣的是，HP1 含一个 chromo 结构域，一种可以结合甲基化 H3 第 9 位赖氨酸的蛋白模块（见第 3 和 4 章）。因此，H2AZ 也许和 H3 甲基化协同作用，为染色质结合蛋白提供一个结合平台。除了其多出来的酸性区域，H2AZ 还有在结构上和一个金属离子相配的一对组氨酸残基。这在体内可能提供了一种 H2A 核小体所没有的生理反应。最后，晶体结构预测了一个主要的 H2A・H2B 二聚体和一个变体 H2AZ・H2B 二聚体组成的不对称组蛋白八聚体。含一个 H2A・H2B 二聚体加一个 H2AZ・H2B 二聚体的核小体会在环 1（Loop1）（图 13-2）处发生冲突，看起来这在体内不可能发生。

总之，H2AZ 核小体的这些新特性指出该变体会给染色质带来独特的物理属性。这个预测得到了实验的证实。例如，H2AZ 可能稳定核小体内二聚体与四聚体间的作用，含 H2AZ 的核小体阵列会显示不同的高级折叠和更少的分子间结合（Dryhurst et al. 2004）。因此，H2AZ 可能以至少三种方式调节染色质功能。首先，它毫无疑问改变了其染色质环境的物理属性，进而影响反式作用因子的接触性或活性。其次，如其他组蛋白，氨基端和羧基端结构域的翻译后修饰可能提供特别的接合位点给染色质结合蛋白（第 3 章介绍的所谓反式效应），或提供电荷密度的变化（顺式效应）。最后，它在染色质上受限制和特异的置入可能将特殊的功能定位到特定基因座。

## 9. H2A 变体组装和置换的蛋白复合物

虽然 H2A 变体如何作用等重要问题依然没有解决，但是最近的研究已经阐述了它们装入染色质的基础。第一个突破来自生化方法纯化出的催化 H2AZ・H2B 二聚体装入染色质的复合物（Sarma and Reinberg 2005）。这个叫做 SWR1-C 的多亚基复合物，含有催化亚基蛋白 Swr1，一个 SWI/SNF ATP 依赖性染色质重塑蛋白家族的成员。在体内，SWR1-C 似乎专一于这项任务，因为删除基因 SWR1 的效果类似于删除基因 H2AZ 本身。而且，swr1 缺失突变体中，H2AZ 在特定基因座的倾向性分布完全丢失。在体外，当纯化的 SWR1-C 和核小体序列放在一起，它能通过 ATP 依赖的反应将 H2A・H2B 替换为 H2AZ・H2B 二聚体（图 13-7）。这个反应有趣的一面来自上文提及的对其结构的一个预测：同时含有 H2A 和 H2AZ 的核小体是不稳定的。因此，SWR1-C 介导的二聚体替换可能是一个协同反应，其中一个 H2AZ・H2B 二聚体的替代促进剩下的 H2A・H2B 二聚体的弹出和替换。

另一个多亚基蛋白复合物是在果蝇中负责把磷酸化 H2AX 替换为未磷酸化分子的（Morrison and Shen 2005）。值得一提的是，这个叫做 dTip60 的果蝇复合物，含有通常见于两个独立复合物的蛋白：前面描述的 ATP 依赖性染色质重塑复合物 SWR1-C 和一个带有乙酰化酶功能的组蛋白修饰复合物 NuA4/Tip60。因此，这一个复合物整合了组蛋白乙酰化、核小体重塑和组蛋白变体替换。这个组合可能反应了果蝇 H2AX 也是它的 H2AZ，而其他真核生物的 H2AX 则是从常规 H2A 进化来的。尽管有这种差异，有理由相信基本通路是保守的。酿酒酵母和哺乳动物细胞中，SWR1-C、NuA4/Tip60 和另一个 ATP 依赖性核小体重塑复合物 INO80-C 共享常见的亚基。其中之一是肌动蛋白相关蛋白 Arp4。有趣的是，Arp4 已知在酿酒酵母中和磷酸化 H2AX 作用，引起后续的 NuA4、SWR1 和 INO80 复合物的招募（Downs et al. 2004）。这暗示了这些复合物催化该物种中 H2AX 和 H2AZ 的置换。这一预测还有待直接证明。

染色质重塑复合物专一于复制非依赖性核小体组装，这一发现的重要性不仅体现在理解组蛋白变体如何装入上，还在于其为染色质重塑装置提供了第一个特异的体内功能。在这些发现以前，我们并不清楚为什么细胞拥有如此大量的大分子机器来促进核小体的运动（Becker and Horz 2002）。SWI/SNF ATP 酶的多样性像一个谜，如果知道一些重塑机器专一于组装不同的变体进入核小体或许能更好的加以理解。也许核小体组装

是一个协调的过程，其中组蛋白修饰酶作用于其底物，而 ATP 依赖染色质重塑蛋白提供复制非依赖的替换过程所需的能量和特异性。

## 10. 其他功能未知的使染色质具有差异的 H2A 变体

脊椎动物中发生了 H2A 的进一步多样化。哺乳动物中，macroH2A 和 $H2A^{Bbd}$（H2A 巴氏小体缺失）代表独特的属系，似乎在剂量补偿的表观遗传现象中起作用（详见第 17 章）。macroH2A 之所以有此称呼是因为除了组蛋白折叠结构域和氨基端及羧基端尾巴，它还含有一个大的羧基端球形结构域（Ladurner 2003）。考虑到 H2A 羧基端尾巴从连接体 DNA 附近伸出来，这个球形结构域有可能和连接体、H3 尾巴或连接体蛋白例如 H1 和高泳动性蛋白（high mobility group protein，HMG）相互作用。只是这个相互作用是什么还不清楚，虽然一个有趣的可能性是它具有酶活性。由于这个 200 个氨基酸球形结构域与具有多聚核苷酸和肽水解活性的蛋白的相似性，这种可能性有所增加。另一种可能是，这个球形结构域可能仅仅是转录起始的一个阻碍，这个角色来自体外实验中它对转录因子结合的封闭作用带来的暗示（Sarma and Reinberg 2005）。macroH2A 的组蛋白折叠结构域也有不同的属性，因为并没有染色质重塑蛋白作用其上。这些观察暗示含 macroH2A 的核小体移动性更差因而可能不利于活跃的转录。这可能是女性兼性失活 X 染色体中 macroH2A 富集于某些分散区域的原因，这些区域和组成型异染色质区域交错分布（图 13-9a）（Chadwick and Willard 2004）。与 macroH2A 相反，$H2A^{Bbd}$似乎在巴氏小体上无法检测到，或者是广泛分布在核中（图 13-9b）（Chadwick and Willard 2001）。$H2A^{Bbd}$核小体在体外的表现和它促进转录的角色相一致（Sarma and Reinberg 2005）。与其他已知 H2A 亚型相比，$H2A^{Bbd}$进化迅速，但是这种快速进化的原因还不清楚。

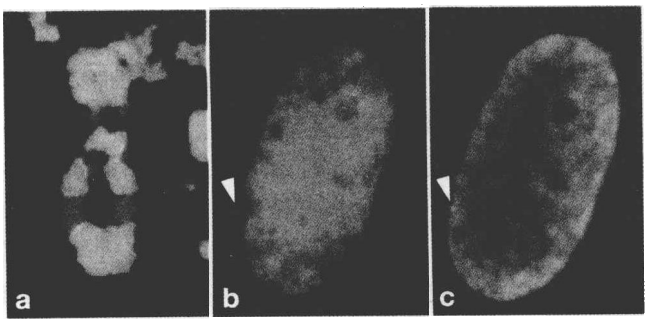



图 13-9　H2A 变体和女性中失活的 X 染色体

(a) macroH2A 抗体（红色）染出的失活 X 染色体分散区域和异染色质的一个标记（组蛋白 H3K9me3）交错分布。(b) $H2A^{Bbd}$（绿色）不分布在失活的 X 染色体上（箭头所指红点）。(c) 与 b 同一细胞核，经 DAPI 染色显示染色质。（a，经许可重印自 Chadwick and Willard 2004[©国家科学院]；b，c，经许可重印自 Chadwick and Willard 2001[©洛克菲勒大学出版社]）。

## 11. 许多组蛋白进化以更紧密地包装DNA

当没有必要为了复制和转录接触DNA的时候，染色质通常会进一步浓缩，二者经常涉及常规组蛋白的替换。显然精子是属于这种情况的，而且在某些属系，组蛋白类似物已经进化出特化的包装功能。例如，海胆精子含有H1和带有结合DNA小沟的重复末端模块的H2B变体（Malik and Henikoff 2003），或许这是为了把染色体包装进入精子头部而产生的适应性。开花植物中花粉特异的H2A变体中可见一种相似的适应性。脊椎动物中，特化的精子特异性组蛋白变体见于哺乳动物睾丸，包括位于精子帽状结构的一个H2B类似体（SubH2Bv）和一个睾丸特异性H3变体（Witt et al. 1996）。

精子成熟过程中组蛋白被替换为精蛋白和其他蛋白的发生提供了一个可能的方式用于擦除雄性生殖细胞中的表观遗传信息。尽管如此，代间遗传的证据（Rakyan and Whitelaw 2003），特别是缺少DNA甲基化的动物中的，提出了一部分核小体组蛋白逃过了这个转换并且传递了表观遗传信息的可能性。像刚指出的，这就是着丝粒处CENP-A所发生的事情（Palmer et al. 1990），而且有可能一小部分变体如H3.3留在精子中以完成基因表达信息的表观遗传学传递。虽然我们对精子发育中组蛋白替换的进程的认识还很初步，我们期望通过理解CENP-A如何逃过这个转化来了解更多。

已完成分裂并进入分化的体细胞中染色体折叠也会增加。在有些情况，折叠包括连接组蛋白量和质的变化。组蛋白H1和核小体的化学剂量关系决定体内核小体序列的平均间隙（Fan et al. 2003）。另外，染色质中H1的存在促进通常抑制转录的染色质高级结构的形成（Wolffe 1992）。体内的连接组蛋白比核心组蛋白移动性强得多。H2A和H2B的停留时间在几个小时，H3和H4的甚至无法测量，然而H1的停留时间是几分钟（Phair et al. 2004）。因此，连接组蛋白变体的装入不大可能会将染色质分化为某种可遗传的形式。H1变体的角色被认为是改变染色质的总体属性，进而影响其总体折叠（Wolffe 1992）。

H1变体和核心组蛋白一样在RC和RI形式间存在区别（Marzluff et al. 2002）。H1的RC变体形式似乎可以相互替换，这是基于五个RC H1变体中一或两个基因被敲除的小鼠具有正常表型的实验的结果（Fan et al. 2003）。鸟类中，连接体组蛋白H5在红细胞成熟过程中置入，伴随着细胞核的极度浓缩。不分裂的细胞中高度累积的另一个变体是$H1^0$，这是高度不同于常规形式的变体。$H1^0$的过表达与类似的常规形式过表达相比，前者使得染色质更少受核酸酶影响。$H1^0$在非分裂细胞中的自然积累可能是细胞活动静止时染色质浓缩的普遍机理。

## 12. 结束语和未来研究

组蛋白变体提供了染色质分化的最基础水平，而置入不同变体的各种机制可能也建立和维持表观遗传状态。组蛋白H2A、H2B、H3和H4占据核心部分上不同的位置，这是在最后一个真核生物共有祖先之前已经发生的一个进化进程的结果。包括从祖先的

类 H3 • H4 四聚体生出八聚体等重要的进化上的革新还是不明确知道的，而我们希望通过测出更多古细菌和早期真核生物的基因组序列以补上其中丢失的环节。紧接着的四个核心组蛋白向不同变体的进化提供了表观遗传学过程的基础，包括发育和染色体分离。为了充分地理解表观遗传信息的继承，我们需要更好的理解替换常规组蛋白而装入变体的进程。一个重要的新近发展是对专一于特定变体的复制非依赖性组装通路的初步描述。

着丝粒是因为一个组蛋白变体的特殊属性而成为显著不同的染色质的最明显例子。虽然很明显 CenH3 核小体是形成着丝粒的基础，但是它们如何在细胞代系传递间没有任何序列特异性的前提下仍置入于相同位置，是将来研究的一个主要挑战。

越来越明显，组蛋白变体也涉及活跃基因的表观遗传属性。H3.3 和 H2AZ 都富集在转录活跃的基因座，而理解造成这种富集的组装过程是当前研究一个激动人心的领域。染色质的动态行为使我们认识到转录、染色质重塑和组蛋白修饰可能耦合于核小体组装和去组装。动态进程搭配组蛋白代谢的研究限于发育早期，而我们分子生物学、细胞遗传学、生物化学和结构生物学的技术进步使我们可以更好的理解染色质的动态属性。

除了这些普遍的进程，组蛋白变体还涉及特殊的表观遗传现象。在哺乳动物 X 染色体中，Y-H2AX、macroH2A 和 $H2A^{Bbd}$ 等三个 H2A 变体能被募集来沉默或激活基因以达到生殖细胞失活或剂量补偿的目的。理解这些变体在表观遗传进程中的功能依然是将来的一个主要挑战。

核小体核心部分的第一个高分辨率结构（Luger et al. 1997）是描述染色质属性的一个重大进步。通过以具有生物意义的方式解开这个基本结构，组蛋白变体提供了一个机会拓深我们对这些奇妙的结构蛋白如何进化以完成表观遗传过程中不同的角色的理解。

谢精卫　译
袁　文　校
朱　冰　审校

## 参 考 文 献


Altheim B.A. and Schultz M.C. 1999. Histone modification governs the cell cycle regulation of a replication-independent chromatin assembly pathway in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci.* **96**: 1345–1350.

Amor D.J., Kalitsis P., Sumer H., and Choo K.H. 2004a. Building the centromere: From foundation proteins to 3D organization. *Trends Cell Biol.* **14**: 359.

Amor D.J., Bentley K., Ryan J., Perry J., Wong L., Slater H., and Choo K.H. 2004b. Human centromere repositioning "in progress". *Proc. Natl. Acad. Sci.* **101**: 6542–6547.

Annunziato A.T. 2005. Split decision: What happens to nucleosomes during DNA replication? *J. Biol. Chem.* **280**: 12065–12068.

Arents G., Burlingame R.W., Wang B.C., Love W.E., and Moudrianakis E.N. 1991. The nucleosomal core histone octamer at 3.1 A resolution: A tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci.* **88**: 10148–10152.

Becker P.B. and Horz W. 2002. ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.* **71**: 247–273.

Belotserkovskaya R., Oh S., Bondarenko V.A., Orphanides G., Studitsky V.M., and Reinberg D. 2003. FACT facilitates transcription-dependent nucleosome alteration. *Science* **301**: 1090–1093.

Chadwick B.P. and Willard H.F. 2001. A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. *J. Cell Biol.* **152**: 375–384.

———. 2004. Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proc. Natl. Acad. Sci.* **101**: 17450–17455.

Downs J.A., Allard S., Jobin-Robitaille O., Javaheri A., Auger A., Bouchard N., Kron S.J., Jackson S.P., and Cote J. 2004. Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Mol. Cell* **16**: 979–990.

Dryhurst D., Thambirajah A.A., and Ausio J. 2004. New twists on H2A.Z: A histone variant with a controversial structural and functional past. *Biochem. Cell Biol.* **82**: 490–497.

Fan Y., Nikitina T., Morin-Kensicki E.M., Zhao J., Magnuson T.R., Woodcock C.L., and Skoultchi A.I. 2003. H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. *Mol. Cell. Biol.* **23**: 4559–4572.

Fernandez-Capetillo O., Mahadevaiah S.K., Celeste A., Romanienko P.J., Camerini-Otero R.D., Bonner W.M., Manova K., Burgoyne P., and Nussenzweig A. 2003. H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. *Dev. Cell* **4**: 497–508.

Formosa T., Ruone S., Adams M.D., Olsen A.E., Eriksson P., Yu Y., Roades A.R., Kaufman P.D., and Stillman D.J. 2002. Defects in SPT16 or POB3 (yFACT) in *Saccharomyces cerevisiae* cause dependence on the Hir/Hpc pathway: Polymerase passage may degrade chromatin structure. *Genetics* **162**: 1557–1571.

Henikoff S. and Ahmad K. 2005. Assembly of variant histones into chromatin. *Annu. Rev. Cell Dev. Biol.* **21**: 133–153.

Jenuwein T. 2001. Re-SET-ting heterochromatin by histone methyltransferases. *Trends Cell Biol.* **11**: 266–273.

Jenuwein T. and Allis C.D. 2001. Translating the histone code. *Science* **293**: 1074–1080.

Kamakaka R.T. and Biggins S. 2005. Histone variants: Deviants? *Genes Dev.* **19**: 295–310.

Krogan N.J., Baetz K., Keogh M.C., Datta N., Sawa C., Kwok T.C., Thompson N.J., Davey M.G., Pootoolal J., Hughes T.R., et al. 2004. Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. *Proc. Natl. Acad. Sci.* **101**: 13513–13518.

Kurdistani S.K., Tavazoie S., and Grunstein M. 2004. Mapping global histone acetylation patterns to gene expression. *Cell* **117**: 721–733.

Ladurner A.G. 2003. Inactivating chromosomes: A macro domain that minimizes transcription. *Mol. Cell* **12**: 1–3.

Lowndes N.F. and Toh G.W. 2005. DNA repair: The importance of phosphorylating histone H2AX. *Curr. Biol.* **15**: R99–R102.

Loyola A. and Almouzni G. 2004. Histone chaperones, a supporting role in the limelight. *Biochim. Biophys. Acta* **1677**: 3–11.

Luger K., Mader A.W., Richmond R.K., Sargent D.F., and Richmond T.J. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260.

Malik H.S. and Henikoff S. 2003. Phylogenomics of the nucleosome. *Nat. Struct. Biol.* **10**: 882–891.

Marc F., Sandman K., Lurz R., and Reeve J.N. 2002. Archaeal histone tetramerization determines DNA affinity and the direction of DNA supercoiling. *J. Biol. Chem.* **277**: 30879–30886.

Marzluff W.F. and Duronio R.J. 2002. Histone mRNA expression: Multiple levels of cell cycle regulation and important developmental consequences. *Curr. Opin. Cell Biol.* **14**: 692–699.

Marzluff W.F., Gongidi P., Woods K.R., Jin J., and Maltais L.J. 2002. The human and mouse replication-dependent histone genes. *Genomics* **80**: 487–498.

Morrison A.J. and Shen X. 2005. DNA repair in the context of chromatin. *Cell Cycle* **4**: 568–571.

O'Neill L.P., Randall T.E., Lavender J., Spotswood H.T., Lee J.T., and Turner B.M. 2003. X-linked genes in female embryonic stem cells carry an epigenetic mark prior to the onset of X inactivation. *Hum. Mol. Genet.* **12**: 1783–1790.

Palmer D.K., O'Day K., and Margolis R.L. 1990. The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. *Chromosoma* **100**: 32–36.

Palmer D.K., O'Day K., Trong H.L., Charbonneau H., and Margolis R.L. 1991. Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc. Natl. Acad. Sci.* **88**: 3734–3738.

Phair R.D., Scaffidi P., Elbi C., Vecerova J., Dey A., Ozato K., Brown D.T., Hager G., Bustin M., and Misteli T. 2004. Global nature of dynamic protein-chromatin interactions in vivo: Three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol. Cell. Biol.* **24**: 6393–6402.

Rakyan V. and Whitelaw E. 2003. Transgenerational epigenetic inheritance. *Curr. Biol.* **13**: R6.

Sarma K. and Reinberg D. 2005. Histone variants meet their match. *Nat. Rev. Mol. Cell Biol.* **6**: 139–149.

Schubeler D., MacAlpine D.M., Scalzo D., Wirbelauer C., Kooperberg C., van Leeuwen F., Gottschling D.E., O'Neill L.P., Turner B.M., Delrow J., et al. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **18**: 1263–1271.

Schwartz B.E. and Ahmad K. 2005. Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev.* **19**: 804–814.

Shi Y., Lan F., Matson C., Mulligan P., Whetstine J.R., Cole P.A., Casero R.A., and Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**: 941–953.

Sullivan B.A. and Karpen G.H. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* **11**: 1076–1083.

Suto R.K., Clarkson M.J., Tremethick D.J., and Luger K. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* **7**: 1121–1124.

Talbert P.B., Bryson T.D., and Henikoff S. 2004. Adaptive evolution of centromere proteins in plants and animals. *J. Biol.* **3**: 18.

Waterborg J.H. 1993. Dynamic methylation of alfalfa histone H3. *J. Biol. Chem.* **268**: 4918–4921.

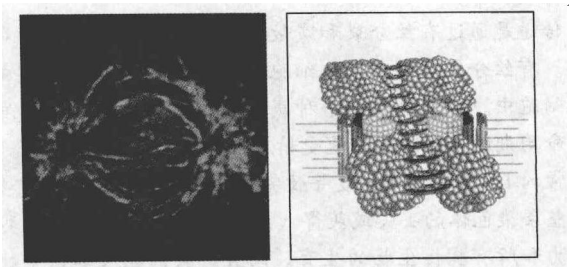
Weiler K.S. and Wakimoto B.T. 1995. Heterochromatin and gene expression in *Drosophila*. *Annu. Rev. Genet.* **29**: 577–605.

Wieland G., Orthaus S., Ohndorf S., Diekmann S., and Hemmerich P. 2004. Functional complementation of human centromere protein A (CENP-A) by Cse4p from *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **24**: 6620–6630.

Witt O., Albig W., and Doenecke D. 1996. Testis-specific expression of a novel human H3 histone gene. *Exp. Cell Res.* **229**: 301–306.

Wolffe A.P. 1992. *Chromatin: Structure and function*. Academic Press, San Diego.

# 第 14 章



# 染色体遗传的表观遗传学调控

Gary H. Karpen[1] and R. Scott Hawley[2]

[1]*Lawrence Berkeley National Laboratory, Department of Genome Biology, University of California at Berkeley, Department of Molecular and Cell Biology, Berkeley, California 94720*

[2]*Stowers Institute for Medical Research, Kansas City, Missouri 64110 and Department of Physiology, University of Kansas Medical Center, Kansas City, Missouri 66160*

## 本章目录

1. 引言
   1.1 染色体遗传是如何完成的?
   1.2 染色体遗传需要哪些元件?
2. DNA复制、修复与端粒的表观遗传调控
   2.1 DNA复制的起始受到表观遗传机制调控
   2.2 DNA修复涉及染色质结构的表观遗传改变
   2.3 端粒结构和功能的表观遗传调控
3. 着丝粒特征和功能的表观遗传调控
   3.1 不同真核生物着丝粒的结构和功能
   3.2 着丝粒序列对于动粒的形成和功能既非必要亦非充分
   3.3 着丝粒染色质的独特构成
   3.4 着丝粒结构，功能和传递模型
   3.5 表观遗传学与着丝粒的演变
4. 异染色质和减数分裂配对/分离
   4.1 雄性果蝇的一个异染色质配对位点的发现
   4.2 雌性果蝇异染色质配对促进分离
   4.3 着丝粒在芽殖酵母无交叉分离中的促进作用
   4.4 玉米中的异染色质相关 *Phi* 基因座及它在介导同源对同源染色体配对中的作用
5. 异染色质和减数分裂驱动
   5.1 雄性果蝇分离紊乱因子
   5.2 蕈蚊中父源染色体丢失和响应元件定位
   5.3 丽蝇金小蜂中父源染色体丢失
   5.4 玉米中的异染色质 knob 10 序列在促进减数分裂Ⅰ染色体分离中的作用
6. 减数分裂过程中非配对DNA导致的基因沉默
   6.1 链孢霉减数分裂中非配对DNA导致的减数分裂沉默
   6.2 小鼠中非联会染色体沉默
   6.3 果蝇性染色体功能紊乱
7. 展望与结论

参考文献

## 概要

遗传信息的复制与传递是通过有丝分裂和减数分裂形式完成的。两者对于生命的生存和进化来说均为最基本的需求。有丝分裂是发生在体细胞的核分裂过程，复制完成的遗传物质以染色体的形式同等地分配到子细胞中。减数分裂是一种减少性的核分裂方式，只发生在多细胞生物体的生殖细胞或在单细胞生命周期的特殊阶段，是受精（或一些真核生物的接合）前产生具有单倍体基因组的细胞的过程。异常的DNA复制或修复导致突变或染色体重排。更重要的是，在核分裂过程中的染色体不分离导致整条染色体的丢失或获得（染色体非整倍性）。这些类型的“基因组不稳定”影响细胞的生存发育能力和所有真核生物的生育。而且，它们在人类出生缺陷和癌症的病因中扮演着关键性的角色。

早期的研究表明DNA序列在减数分裂和有丝分裂所需染色体元件位置和功能确定方面起到主导作用，如DNA复制的起始点，纺锤体接触位点（动粒和着丝粒），染色体的终止点（端粒）和减数分裂配对位点的确定。然而，在最近的十年里我们发现表观遗传机制调控着基因组稳定和染色体遗传所需的许多关键功能，这其中就包括DNA复制起始点的确定、DNA修复和重组、染色体的保护（端粒）、染色体移动（着丝粒）和减数分裂过程中同源染色体分离。乍看表观遗传调控似乎矛盾于染色体功能是细胞与生物体生存所必需的事实，事实却暗示染色体在DNA序列上是固定不变的“硬件。然而，从进化的角度观察，染色体在减数分裂和有丝分裂过程所显示的表观遗传的“可塑性”似乎是对物种形成相关序列改变和染色体重排的重要补充。理解表观遗传学调控的分子基础对于阐明这些基本的生物学过程以及人类疾病的诊断和治疗具有重要的作用。

# 1. 引　言

## 1.1　染色体遗传是如何完成的?

有丝分裂是一种基本的细胞分裂方式，产生等同的二倍体子细胞，通常发生在体细胞或减数分裂前的生殖细胞中（图 14-1a）。细胞有丝分裂周期包括四个阶段，称为 $G_1$ 期（gap1，有丝分裂后的“休息”时期），S 期（synthesis，合成期，DNA 复制和基因表达），$G_2$ 期（gap2，S 期后的“休息”阶段，为有丝分裂做准备）和 M 期（mitosis，有丝分裂期，包括前期、中期、后期和末期）。在 S 期 DNA 发生复制，复制形成的姐妹染色单体连接在一起。在有丝分裂开始时，染色体浓缩，组蛋白 H3 第 10 位丝氨酸磷酸化（H3S10ph）（见第 10 章）。另外，在大多数的有机体中，在每一个姐妹染色单体上的一个单一位点（动粒，kinetochore）形成一个称为着丝粒（centromere）的结构，其介导与纺锤体微管的接触并作为一个细胞周期检验点。成对的姐妹染色单体在前中期聚集到赤道板，在后期分开移向两极。这些移动是由着丝粒结合的微管运动蛋白（驱动蛋白和动力蛋白）和微管的装配、去装配调控的，也需要从中期到后期的转化过程中姐妹染色单体结合位点的破坏。

减数分裂只发生在生殖细胞并且其特点是一次复制伴随着两次分裂（减数分裂Ⅰ和Ⅱ，图 14-1b）；生成了多细胞动物雌性的单倍体卵细胞和雄性的单倍体精细胞，而非二倍体的子细胞。在减数分裂Ⅰ期，复制的同源染色体相互配对并同时分离。直到减数分裂Ⅱ期每个同源的姐妹染色单体才彼此分开。减数分裂过程中的正常分离需要同源染色体之间频繁的发生重组，在动粒区域也发生特异性的附着以确保姐妹染色单体在减数分裂Ⅰ期的结合（Watanabe 2005）。

## 1.2　染色体遗传需要哪些元件?

细胞有丝分裂和减数分裂都需要特定的染色体元件的活性和结合蛋白以完成精确的基因组复制和染色体分离（图 14-2）。DNA 的复制是在“起始点”开始的，在多数真核生物这一过程并不是严格的序列依赖性的（讨论见第 2 节）。在有丝分裂的细胞中姐妹染色单体的结合在整个染色单体上都逐渐可见，尽管在着丝粒周边异染色质存在高度浓缩的黏结蛋白。动粒是一个大的区域，由 DNA 和特异的染色质蛋白组成，作为着丝粒形成的基础和纺锤体结合以及正常有丝分裂和减数分裂染色体分离的关键。在多数真核生物中，每个染色体有且只有一个动粒。如果动粒缺失会导致纺锤体结合的失败和染色体丢失，而若出现多于一个动粒会导致同一染色单体于两极进行连接，会引起后期染色体桥联和断裂。少数情况下，生物体（如 *C. elegans*）具有多着丝粒（polycentric）或全着丝粒（holocentric）染色体，在其上着丝粒在多个区域出现（如图 13-5）。这些染色体通过特殊的机制确保姐妹染色单体接触和向相对的两极移动。端粒是特殊的染色质结构，位于染色体末端能够保护其不受降解或重组，并且能确保 DNA 的复制完全。减数分裂分离也需要着丝粒、端粒、黏附分子和复制起点。然而，需要更多其他的元件

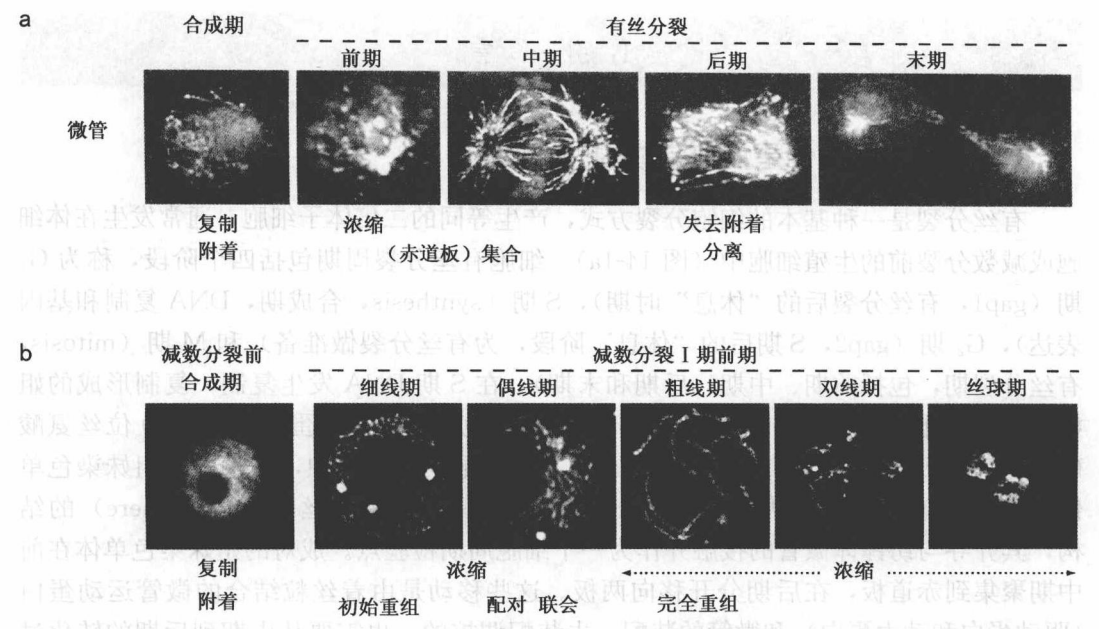



图 14-1 有丝分裂和减数分裂各阶段

(a) 来自果蝇细胞的图示描述染色体（蓝色）、微管（绿色）和着丝粒（红色）在分裂间期和有丝分裂期的行为。(b) 玉米减数分裂Ⅰ期前期的染色体行为，在此阶段同源染色体发生配对、联会和重组（照片由佛罗里达州立大学的 Hank Bass and Shaun Murphy 提供）。在各个阶段关键性的染色体功能描述见下方蓝色字体。随后，在减数分裂Ⅰ期的后期同源染色体分离移向两极，完成一次减数分裂。姐妹染色单体只在减数分裂Ⅱ期发生分离（图 14-9）。

和着丝粒行为修饰来确保减数分裂Ⅰ期的同源体配对和分离（讨论见第 4 节）。

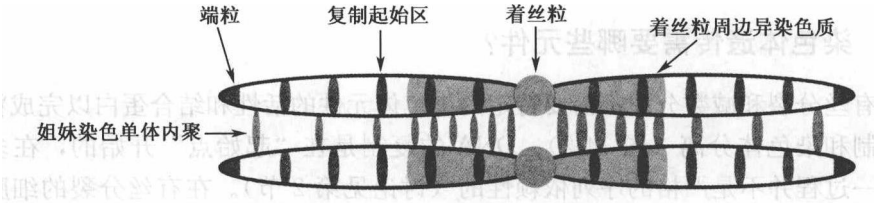



图 14-2 染色体遗传元件

有丝分裂和减数分裂中对于正常复制（复制起始点）和遗传（着丝粒，黏附蛋白，端粒）所必需的染色体元件。正常减数分裂的分离和多数情况下的重组都需要同源体配对位点（图中未标出）。

染色体遗传的关键特性表明遗传元件的特异性和定位应该是 DNA 序列的“硬线”(hard-wired)* 。然而，令我们吃惊的是，许多元件，包括在这一章节中提到的几个，却是受到表观遗传的调控，尤其是多细胞真核生物。总之，那些倾向于受表观遗传调控的元件确保了染色体遗传的保真性，包括 DNA 复制起始点、端粒、姐妹染色单体黏附

* 意为 DNA 序列是决定性因素——译者

位点、着丝粒和同源体配对位点。

## 2. DNA 复制、修复与端粒的表观遗传调控

确保遗传信息正确传递的第一步包括整个基因组副本忠实地复制，这是通过 DNA 复制完成的。不幸的是，复制过程会出现错误，导致 DNA 的改变（突变），这会对生命体的生存造成威胁。而且，环境的因素（如辐射）能够产生突变，包括 DNA 碱基的改变、缺失、插入和重排。细胞通过激活 DNA 修复途径来对 DNA 的损伤发生响应，这些途径的作用令人惊叹地保证了基因组的保真性和稳定性。最后，线性 DNA 分子的复制遇到的挑战被位于染色体末端的特殊序列和结构，即端粒所克服。最近的研究表明，DNA 准确复制和维持所需的基本过程都受到染色质的表观遗传学调控机制的影响。

### 2.1　DNA 复制的起始受到表观遗传机制调控

DNA 忠实的复制是由具有 5′到 3′活性的 DNA 聚合酶完成的，在称为“起始点”的特异位点起始。一个“复制泡”包括两个在起始点形成“复制叉”，复制是双向进行的，直到与另一个复制起始点产生的复制叉相遇。由单一的复制起始点起始复制的区域通常非常大，覆盖上百 kb（Aladjem and Fanning 2004）。酿酒酵母的自主复制序列（autonomously replication sequence，ARS）克隆到质粒中后发生功能异常，并通常整合入其他的染色体位点，表明复制的起始是受特异 DNA 序列调节的。ARS 约 100～150bp，包含一个或多个约 11bp 的富含 AT 的重复的关键序列，还包括其他一些保守序列（Weinreich et al. 2004）。一个称为复制起始点识别复合体的蛋白复合物（origin recognition complex，ORC）是起始必需的并且负责募集复制前复合物（prereplication complex，PRC），包括小染色体维持蛋白（minichromosome maintenance，MCM）（Prasanth et al. 2004）。对于多细胞动物，复制起始点能够被原位检测到，但是它们被克隆或再引入时都表现为特征性的失活状态，表明这些生物的复制起始是受到表观遗传调控的而非严格的 DNA 序列依赖的（Aladjem and Fanning 2004）。虽然 ORC 和 MCM 蛋白在多细胞生物中是保守的，但是这些生物中与调节起始活性相关的因子和机制仍然未知。而且，现在我们仍然不清楚染色质的结构是否影响复制叉的连续性。

有关多细胞生物复制起始可能受表观遗传调控的线索来源于对酿酒酵母的详细研究。虽然复制起始点的 DNA 序列是必需的而且它已经足够发动起始过程，但仍有明确的证据表明染色质结构对起始活性有影响（Weinreich et al. 2004）。基因芯片研究表明，并不是所有的 322 个双向复制位点和 429 个 ORC 或 MCM 蛋白结合位点在每一个细胞周期中都有活性。染色体所处环境和染色质结构能影响假定的起始点在复制起始时是否被活化。例如，一个沉默接合型位点的 ARS 不能起始复制除非移动到染色体的其他位点。另一个起始的表观遗传调控的例子包括约 100～200 个基因编码的核糖体 RNA（ribosomal RNA，rDNA），它是串联重复排列的（Pasero et al. 2002）。每 9.1kb rDNA 单位包含一个 ARS，在被插入到质粒中或基因组其他位置时都能够起始复制。然而，只有约 20%的 rDNA 起始点在 S 期具有活性。

起始点也能够在S期不同的阶段被激活（Weinreich et al. 2004）。例如，靠近端粒的起始点正常情况下是在S期后期激活（见第2.3节），但是把相同的序列插入到环形质粒中则会在S期的早期起始复制（图14-3）（Ferguson and Fangman 1992）。晚期复制概括起来主要是质粒线性化或端粒的附加，后者使ARS临近染色体末端。

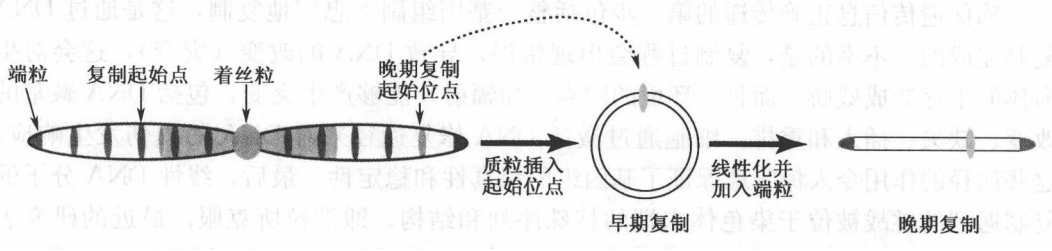



图14-3 酵母复制时间的表观遗传调控

酿酒酵母是表观遗传影响复制的最好的例子。晚期复制起始位点插入到环形质粒中能引起S期的早期复制，而质粒线性化和端粒的加入能够恢复晚期复制（Ferguson and Fangman 1992；Weinreich er al. 2004）。

这些结果表明，复制起始点的活性和在S期的激活时间都是受表观遗传调控的。进一步的证据来自于S期时序发生和rDNA、端粒以及接合型位点起始沉默都是受许多组蛋白修饰和表观遗传基因沉默相关蛋白（如SIR）调节的（见第4章）。

## 2.2 DNA修复涉及染色质结构的表观遗传改变

由复制错误或环境因素引起的DNA损伤积累，可以导致有害的突变、基因组不稳定、癌症、细胞衰老和死亡。在复制过程中DNA损伤能够被纠错机制修复，也能够通过独立的通路$G_2$期进行修复。细胞拥有“检验点”能够识别DNA损伤的存在并停滞或延迟细胞周期直到修复完成；如果损伤不能被修复，这些通路也能诱导细胞死亡（凋亡），由此通过移除缺陷细胞而维持生物的生存。这些途径通常是非常有效的；例如，人皮肤细胞暴露在阳光的UV射线后会含有大量的损伤，其中的大部分能够被正确的修复（Friedberg et al. 1995）。有修复缺陷的个体，可能是因为检验点或修复途径的一个或多个成分突变，会受到许多疾病的困扰，包括易患结肠、乳腺和皮肤癌症以及早老。

早期研究成功地鉴别了能够识别DNA分子中不同类型损伤的分子和通路，如双链断裂（double-strand break，DSB）、嘧啶二聚体，还发现了能够募集特异复合物修复损伤的物质。DNA包装成染色质可能潜在地阻碍了因子对损伤位点的识别或影响了修复，这类似于异染色质对基因表达的抑制作用（Hassa and Hottiger 2005）。然而，修复有赖于ATP依赖的染色质重塑复合物，推测其可能是把缺陷DNA暴露以便进行修复。

最近的研究显示染色质模板的特异性改变（如组蛋白变体的出现和转录后组蛋白修饰），在DNA损伤修复的识别和修复复合体的募集中起关键作用（Hassa and Hottiger 2005）。例如，DSB能引起组蛋白H2A变体H2AX的136和139位丝氨酸的快速磷酸化（称为γH2AX）。H2AX磷酸化对于修复蛋白在环绕DSB的大范围（Mb级）的累积是必需的，对修复区域的装配也是需要的，而对于DNA主要损伤位点的修复复合物

的初始募集并不必需（Bassing et al. 2002；Celeste et al. 2002）。这些发现表明，从 DSB 蔓延开的 γH2AX 的活化将扩大来自 DSB 的信号，加强募集或者能够固定修复因子（Fernandez-Capetillo et al. 2004）。除了在 DSB 修复时涉及之外，H2AX 磷酸化影响 V（D）J 序列在哺乳动物淋巴细胞中的重排并可作为基因组不稳定性和小鼠肿瘤发生的抑制因子发挥作用（Fernandez-Capetillo et al. 2004）。

其他形式的染色质改变，如组蛋白乙酰化、SUMO 化、泛素化和甲基化也均有证据表明在 DNA 损伤的成功修复中起到一定作用。例如，组蛋白 H3 的 79 位赖氨酸甲基化（H3K79me）对于修复检验点蛋白 53BP1 在 DSB 的募集是需要的，并且由 DOT1 组蛋白甲基转移酶（histone lysine methyltransferase，HKMT）介导（Huyen et al. 2004）。有意思的是，诱导的 DNA 损伤并不改变 H3K79 的甲基化水平，表明这一修饰并不随对 DSB 的反应而增加。对此一种可能的解释是 53BP1 的募集和 DSB 的“感知”影响到已存在的 H3K79 甲基化的暴露并回应为 DNA 损伤位点的染色质的重塑。

虽然我们现在对染色质因子对 DNA 损伤的影响的理解表明它们在信号转导和适当的复合体在 DNA 损伤部位的募集中的作用，但今后的研究可能会揭示更多的表观遗传机制调节基因组稳定性的方式。有一点是很重要的，即染色质在 DNA 修复中的作用是主动的和因损伤而发生的。虽然组蛋白修饰和其他的表观遗传调节蛋白发挥了关键作用，但这些改变在细胞分裂过程中并不能被继承，这与本章中讨论的其他例子有所不同。

### 2.3　端粒结构和功能的表观遗传调控

线性真核生物染色体的末端是特异化的结构，称为端粒，它具有三个关键作用：①确保包括染色体最末端在内的 DNA 复制，解决了“末端复制问题”（Lue 2004）；②防止染色体末端被降解和与其他染色体的融合；③在许多但不是所有的有机体中，端粒有益于染色体在减数分裂中的配对。在多数真核生物中，端粒由简单的短重复序列组成，是由端粒酶修复和重建的。端粒的功能受序列的影响并有表观遗传机制共同参与调节。

末端复制问题的产生是由于 DNA 聚合酶需要一个引物才能起始 5′到 3′后随链的合成；这样，活性受限制致使复制不能一直进行到染色体的末端（Lue 2004）。有两种机制能够解决这一问题，其中主要的被多数有机体（包括酵母、哺乳动物和植物等）所采用的机制是一种被称之为端粒酶的特殊酶复合物。端粒酶在多数真核生物中是由简单的 6bp 的重复序列延伸成数十或数百个 kb。端粒酶复合体含有类逆转录酶活性及与端粒重复序列有同源性的 RNA 片段。本质上末端的复制完成包括通过 RNA 成分把复合物定位至端粒重复序列，接着就是逆转录（3′到 5′）产生新的重复序列。有意思的是，端粒酶活性的缺失或端粒的缩短与细胞的衰老与老化有关；相反，癌细胞具有增强的端粒酶活性和延长的端粒（Blasco 2005）。

端粒酶非依赖的染色体末端维护机制也存在（Louis and Vershinin 2005）。其中一个研究较透彻的系统是在果蝇和其他双翅目昆虫中。这类生物缺少明确的端粒酶，并且末端也没有其他真核生物中发现的简单重复序列。取代它的是染色体末端成簇散乱的非

长末端重复片段（long terminal repeat，LTR）逆转座子，其大小为 3～5kb 不等，除此之外还有其他的重复序列，如端粒结合重复序列（telomere-associated repeats，TAS）（Biessmann and Mason 2003）。这些转座子编码逆转录酶（因此称为逆转座子），表明可能在进化上与更标准的端粒酶机制有一定关系。然而，主要的不同是果蝇的染色体并复制到最末端；每一代果蝇将丢失 70bp 左右，这一数值是由末端复制问题粗略估计的。这一损失并不引起关键基因的缺失，因为端粒和亚端粒重复区域长达 50～100kb，要经过很多代才会损失足够的 DNA 到达基因区域。然而，端粒序列的丢失能够通过非经常性的非 LTR 逆转座子的加入而得以补偿（Biessmann and Mason 2003）。

表观遗传调控影响着端粒的功能以及周围基因的表达。“裸露”的端粒 DNA 和内含的 DSB 都会引起染色体的融合和非整倍性。Barbara McClintock 首先描述了一种称为断裂-融合-桥周期（breakage-fusion-bridge cycle）的现象，断裂的染色体或染色体末端融合，产生具有双着丝粒的染色体和分裂后期的染色体桥，这又能产生进一步的断裂。端粒末端保护作用的表观遗传学调节的证据来自果蝇的研究，表明这一调节是 DNA 序列非依赖性的。在果蝇某一代中断裂的染色体的行为类似 DSB，但随后能发挥端粒的全部功能，而且不伴有任何附加的逆转座子和序列改变（Ahmad and Golic 1998）。而且，在果蝇中产生的任何末端能够包装成端粒并能防止发生融合（Karpen and Spradling 1992）。进一步研究发现，裂殖酵母中的端粒功能的发挥依赖 Taz1 蛋白（Miller and Cooper 2003）和端粒染色质，而不依赖于端粒重复序列（Sadaie et al. 2003）。

端粒区域包含染色质修饰和类似第 3 节中描述的着丝粒周边异染色质的特性。表观遗传调节端粒和亚端粒区域的机制来源于对酵母和果蝇中基因表达的研究，但也会在人类中发生。常染色质基因插入到端粒区域后能发生可变性沉默，这被称为端粒位置效应（telomere position effect，TPE）并且与在果蝇和裂殖酵母中由邻近着丝粒异染色质诱导的位置效应花斑相似（position-effect variegation，PEV）（更多的细节请参阅第 5 和 6 章）。在酿酒酵母许多这些特殊的染色质相关因子如 SIR 蛋白，能影响接合型沉默，也能影响端粒诱导的沉默（见第 4 章）。令人吃惊的是，几乎所有在果蝇中已知的能调控 PEV 的基因（位置效应花斑抑制子和增强子，*Su(var)* 和 *E(var)*，参阅第 5 章）没有能够影响端粒沉默的。这表明 PEV 和 TPE 至少是部分通过不同通路介导的（Cryderman et al. 1999；Donaldson et al. 2002）。

异染色质蛋白 1（heterochromatin protein，HP1，一个 *Su(var)* 基因产物）和 H3K9 甲基化，都是异染色质介导的沉默的关键成分（见第 8 章），在果蝇端粒中出现并都是端粒延伸所需（图 14-4）（Perrini et al. 2004）。HP1 或它的结合配基 HOAP（(HP1/ORC associated protein，HP1/ORC 相关蛋白）的去除，会导致很高频率的端粒融合（Cenci et al. 2003）。HP1 通常是经由 Chromo 结构域通过其与甲基化的 H3K9 的亲和力被募集到染色质。有趣的是，端粒的通过 HP1 的加帽反应是不依赖于 H3K9 甲基化的，说明末端保护是通过其他机制介导的包括直接结合到端粒 DNA 或在末端删除中出现的非端粒序列（图 14-4a）（Perrini et al. 2004）。一个理想的模型是：HP1 结合并保护末端非 DNA 依赖序列，然后募集一个未知的 H3K9 HKMT；局部的 H3K9

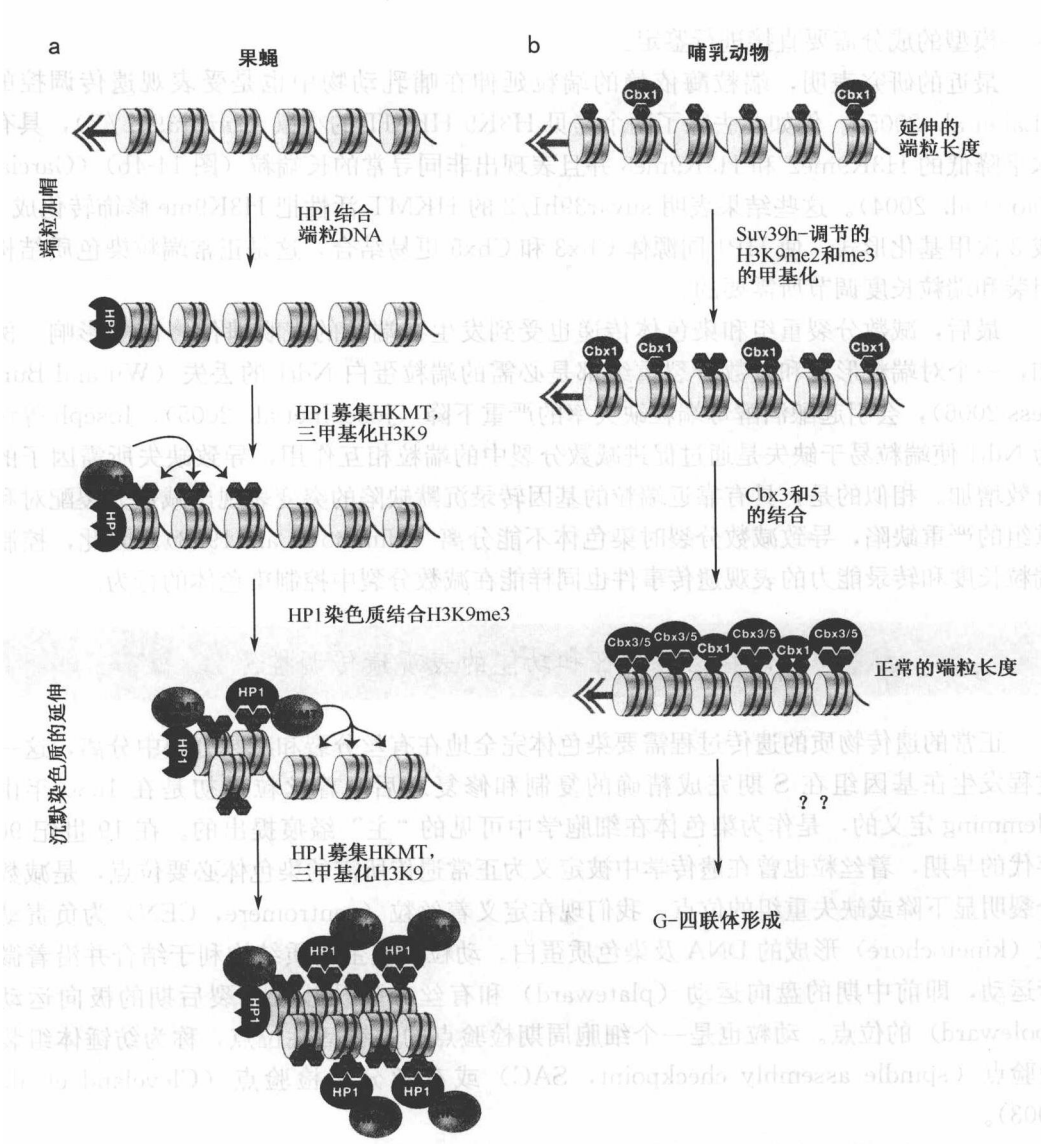



图 14-4　果蝇和哺乳动物中端粒功能的表观遗传调控

(a) 果蝇中，异染色质蛋白 1(HP1) 不依赖 chromo 结构域结合端粒 DNA；端粒加帽，能够通过阻断端粒融合确保正常的分离 (Fanti er al. 1998；Perrini et al. 2004)。然后 HP1 募集一个未知的组蛋白甲基转移酶 (HKMT；不是 *Su(var)*3-9) 在邻近的核小体上三次甲基化 H3K9；HP1 通过 chromo 结构域结合 H3K9me3，这反过来又能募集更多的 HKMT，连续的环形的 HP1 结合 HKMT 募集促进染色质沉默通过亚端粒区域蔓延。(b) 小鼠中，敲除两个 *Suv 39* HKMT 位点导致 H3K9me3 和 me2 的水平降低、H3K9me 修饰水平升高、染色质结构改变，结合位于端粒的双和三甲基化的 H3K9(↓Cbx3 和 5)，H3K9me(↑Cbx1) 和 TERFs1 和 2 (未显示) 蛋白水平的改变 (Garcia-Cao et al. 2004)。这些改变与端粒长度延长有关，表明 *Suv 39hs* 介导的 H3K9 三甲基化对于正常端粒融合与大小的调节是需要的。

甲基化能随后募集更多的 HP1 到此区域，进而促进端粒沉默的延伸 (图 14-4)。这一机制可能不需要在着丝粒异染色质的建立和沉默中涉及的 RNAi 通路 (见第 8 章)，但是

这一模型的成分需要直接进行鉴定。

最近的研究表明，端粒酶依赖的端粒延伸在哺乳动物中也是受表观遗传调控的（Lai et al. 2005）。例如，去除了两个拷贝 H3K9 HKMT 的小鼠（*suvar39h1/2*），具有水平降低的 H3K9me2 和 H3K9me3 并且表现出非同寻常的长端粒（图 14-4b）（Garcia-Cao et al. 2004）。这些结果表明 suvar39h1/2 的 HKMT 活性把 H3K9me 修饰转化成 2 或 3 次甲基化形式，使 HP1 同源体 Cbx3 和 Cbx5 更易结合，这是正常端粒染色质结构组装和端粒长度调节所需要的。

最后，减数分裂重组和染色体传递也受到发生在端粒的表观遗传修饰的影响。例如，一个对端粒形成和减数分裂重组都是必需的端粒蛋白 Ndj1 的丢失（Wu and Burgess 2006），会引起酿酒酵母端粒缺失率的严重下降（Joseph et al. 2005）。Joseph 等认为 Ndj1 使端粒易于缺失是通过促进减数分裂中的端粒相互作用，导致缺失所需因子的有效增加。相似的是，具有靠近端粒的基因转录沉默缺陷的突变表现出减数分裂配对和重组的严重缺陷，导致减数分裂时染色体不能分离（Nimmo et al. 1998）。因此，控制端粒长度和转录能力的表观遗传事件也同样能在减数分裂中控制染色体的行为。

## 3. 着丝粒特征和功能的表观遗传调控

正常的遗传物质的遗传过程需要染色体完全地在有丝分裂和减数分裂中分离，这一过程发生在基因组在 S 期完成精确的复制和修复之后。着丝粒最初是在 1880 年由 Flemming 定义的，是作为染色体在细胞学中可见的“主”缢痕提出的。在 19 世纪 90 年代的早期，着丝粒也曾在遗传学中被定义为正常遗传所需的染色体必要位点，是减数分裂明显下降或缺失重组的位点。我们现在定义着丝粒（centromere，CEN）为负责动粒（kinetochore）形成的 DNA 及染色质蛋白。动粒是使蛋白质结构利于结合并沿着微管运动，即前中期的盘向运动（plateward）和有丝分裂及减数分裂后期的极向运动（poleward）的位点。动粒也是一个细胞周期检验点的关键活性位点，称为纺锤体组装检验点（spindle assembly checkpoint，SAC）或有丝分裂检验点（Cleveland et al. 2003）。

对于单着丝粒染色体生物的关键问题是每一个染色体有一个且仅有一个位点，它是如何与着丝粒功能相联系的（称为“着丝粒特征”，centromere identity）以及这些信息是如何从一个细胞或一代生物传到下一个细胞或下一代生物的（称为“着丝粒传递”，centromere propagation）。这里我们的证据表明在大多数真核生物中着丝粒特征和传递是通过染色质结构而非特异的 DNA 序列来进行表观遗传调控的（Carroll and Straight 2006）。这些数据的关键内容可以总结为：①着丝粒的序列即使在关系紧密的物种间甚至在单一物种的染色体间也不是保守的；②着丝粒 DNA 对于动粒的形成是并非必需也非充分的；③在进化过程中着丝粒在染色体上的位置表现出很大的可塑性。

### 3.1 不同真核生物着丝粒的结构和功能

20 世纪 80 年代对酿酒酵母的研究导致了第一次对真核生物着丝粒的克隆和分析。

一个 125bp 的结构出现在所有 16 个酿酒酵母染色体中，它被证明是对正常着丝粒功能必需且充分的（Bloom et al. 1989）；在这高度保守的元件Ⅰ和Ⅲ中，即使是单个碱基的改变都会导致功能的完全丧失。因此，在单细胞真核生物中着丝粒的特征和传递是由 DNA 序列决定的。

可以调节着丝粒特征的类似的序列在研究另一个简单的真核生物裂殖酵母时破灭了。在裂殖酵母中，着丝粒序列在结构上比酿酒酵母中更大更复杂。（Clarke et al. 1986；Nakaseko et al. 1987）。动粒形成位点的非同源的 4～5kb 长的“中心核”序列，其周边是三个染色体共同具有的各种类型的反向重复序列。最小的一个 25kb，包含非重复中心核，内重复序列和一部分外重复序列，是着丝粒功能和稳定染色体传递所必需的（Baum et al. 1994）。合理的着丝粒功能能够通过转染带有一个中心核和内重复序列（如中心域）及两侧外重复序列的质粒观察到。有趣的是，内重复序列的缺失过程可使减数分裂的姐妹染色单体分离，说明着丝粒区域在动粒组装外的其他过程中也发挥作用。确实，动粒和黏附区域的关系非常紧密而且对于染色体的正确分离非常重要。

虽然着丝粒区域在多细胞真核生物中比裂殖酵母（几百或几千个碱基的重复 DNA 序列）中的更大更复杂，但裂殖酵母着丝粒的总体结构和功能却能够作为研究哺乳动物、植物和昆虫着丝粒的极好模型。着丝粒在这些生物体中嵌入到出现在各染色体上的大块异染色质中，其主要是由卫星 DNA（简单的短重复序列）和转座子组成。这些着丝粒区域是由子域组成，子域负责不同的功能尤其是动粒的形成和姐妹染色体的粘合。然而，着丝粒序列在真核生物中并不是保守的，甚至在同一物种不同染色体中也是不同的。只是着丝粒功能子域的表观遗传组成表现出保守性，显然其是通过组蛋白变体和受表观遗传调控的组蛋白修饰方式来实现的。

秀丽隐杆线虫和其他物种中，散漫着丝粒染色体（holocentric chromosome）沿着整个染色体长度募集和装配着丝粒蛋白（Dernburg 2001）。特异的 DNA 序列是不需要的，正如 λ 噬菌体的串联体和许多其他 DNA 能够被稳定的传递。蛋白在前期成束状募集，但在中期则沿着染色体臂极向面均匀的分布，证明秀丽隐杆线虫基因组的许多区域都支持动粒以表观遗传的方式装配。尽管单着丝粒染色体有着明显的不同，其在组织和结构上的属性（如 CEN DNA 的 3D 螺旋和环）的保守性还是可能的（见 3.3 节）。

## 3.2　着丝粒序列对于动粒的形成和功能既非必要亦非充分

在酵母 DNA 序列分析研究中运用的很成功的各种构建物很难用于多细胞真核生物。因为后者有大而复杂的着丝粒序列。虽然已经通过在组织培养细胞转染一系列卫星 DNA 而产生了少量人类人工染色体（human artificial chromosome，HAC），但其具有有丝分裂的高度不稳定性（Rudd et al. 2003）。然而，我们知道 HAC 是由引入的卫星序列的 DNA 环化形成的，但是一些 alpha 卫星排列则不能够重复形成着丝粒，表明需要更多未知的步骤和因子。最近的研究发现了着丝粒染色质独特的特性和成分（解释见第 3.3 节）。其在 HAC 的卫星排列和非着丝粒序列（如质粒载体和选择性序列标记）上都有发现（Lam et al. 2006）。因此，特异的 DNA 序列在组装和维持功能性人着丝粒上是否足够依然不清楚。

最早的关于着丝粒特征和传递是受表观遗传调控的研究结果来自对裂殖酵母中的最小的着丝粒结构的研究（Steiner and Clarke 1994）。低频率的转化株表现出减低的着丝粒功能和高度激活的着丝粒活性转换（0.6%细胞），随后又能永久地建立于许多的传代中。因此同一 DNA 序列呈现出功能上不同的两种可遗传状态，类似于基因表达上观察到的表观遗传对 PEV（第 5 章）和 TPE（第 4 章）的效应。

其他的观察结果也强烈地支持表观遗传机制在决定多细胞生物着丝粒特征性和动粒形成中的作用。首先，DNA 序列与着丝粒正常的联系对其功能来说是不足够的。例如，只有少部分小鼠和人的异染色质卫星序列与着丝粒功能有关（Lam et al. 2006）。而且，在果蝇和人中发现的具有两个着丝粒卫星区域的功能性染色体（双着丝粒），其中一个区域缺失了形成动粒的能力（Sullivan and Willard 1998）。其次，既然非着丝粒 DNA 能够通过一个称为“新着丝粒形成”的过程获得和传递着丝粒功能，因此着丝粒序列对于动粒的形成来说并不是必需的（图 14-5a）。人类许多功能性的新着丝粒已被鉴定，而且序列分析发现新的动粒形成区域没有获得卫星 DNA。然而，新动粒两侧的序列获得了与内生性着丝粒（即着丝粒周边异染色质）相应区域相似的表观遗传特性，如 H3K9 甲基化和 HP1 的结合（Lo et al. 2001）。

虽然人类中新着丝粒形成的机制还不清楚，但是新着丝粒已经能够在一个模型系统中实验性地产生了。在果蝇中，新着丝粒能够在非着丝粒 DNA 和一个内源着丝粒并列时由小染色体产生（图 14-5b）（Maggert and Karpen 2001）。因此，邻近功能性着丝粒是在果蝇中产生新着丝粒活性必需的，表明获得新着丝粒的一个机制是着丝粒蛋白顺式延伸到邻近非着丝粒区域。一旦这一延伸发生，着丝粒功能就会在这一位置发生表观遗传性传递。有趣的是，当新着丝粒形成区域和内源着丝粒之间出现异染色质时新着丝粒的形成就会受阻，表明还有其他的表观遗传机制参与决定着丝粒大小。

最后，染色体重排是进化和物种形成的一个特点。这些改变伴有着丝粒的获得、丢失和相对于基因组序列的移动（Ferreri et al. 2005）。如果着丝粒特征性是如同 3.5 节中描述的一样由表观遗传决定的，那么这种可塑性就变得好解释了。

## 3.3 着丝粒染色质的独特构成

着丝粒特征和传递的表观遗传调控证据指向一种可能，即染色质的结构和成分是关键的决定因素，而 DNA 序列并非主要。这里，我们讨论着丝粒染色质中发现的独特成分和结构以及一个令人吃惊的发现，即这些特征在远源真核生物中是保守的。

着丝粒特异性组蛋白——H3 样蛋白 CENP-A 家族出现在所有真核生物着丝粒核小体中（图 14-6a）。它们是动粒的结构和功能基础也是很好的建立和传递着丝粒特征的表观遗传标记的候选者（Cleveland et al. 2003）。不像大多数的动粒蛋白在有丝分裂时装配，CENP-A 在整个细胞周期的着丝粒上都能见到，这也是它作为着丝粒特征性的首要表现。CENP-A 含有的染色质也为动粒蛋白募集，纺锤体接触以及酵母、蠕虫、果蝇和哺乳动物的正常染色体分离提供了一个的必要基础（Carroll and Straight 2006）。相互的上位试验表明，CENP-A 是动粒装配通路中的第一个蛋白，与有丝分裂染色体动粒基部染色质共定位。进一步说明 CENP-A 在动粒形成中重要性的证据来自果蝇中

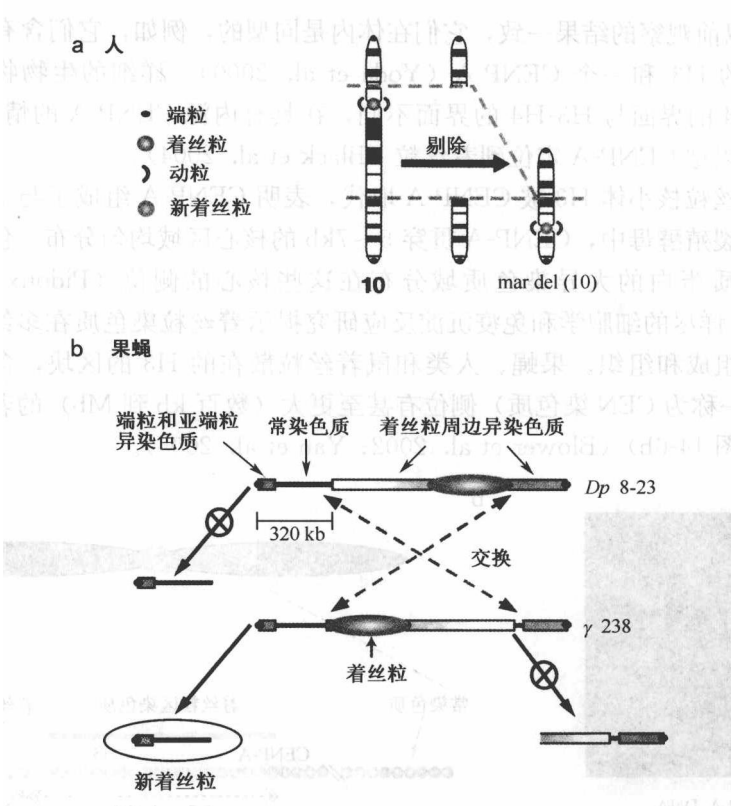



图 14-5　人类和果蝇中新着丝粒形成

(a) 携有新着丝粒的人染色体常常与总 DNA 重排有关，该染色体在着丝粒 DNA 缺失时展现出着丝粒功能并形成动粒（Amor and Choo 2002）。在这一经典范例中，源于新着丝粒的 10 号染色体（mar(del)10）的结构表明，它的形成是通过一个移除了内源性着丝粒（灰点线）的大的 DNA 删除实现的。mar(del)10 在一个包含单个环形染色体（ring(del)10，未标出）核型的个体中恢复，这一环形染色质包含删除区域的 DNA。人新着丝粒形成的顺序依然不清楚；新着丝粒的形成可能先发生，产生一个双着丝粒染色体，随后发生重排，或者新着丝粒也可能发生在内源性着丝粒丢失之后。(b) 在果蝇中，新着丝粒可以试验性地从一个分子水平定义的小染色体产生。一个不含有着丝粒 DNA 的 320kb 的常染色质片段和端粒染色质，能够通过辐射从这一小染色体的剩余部分分离。这一原本“无着丝粒的”片段，能够转变成一个功能性的新着丝粒，在有丝分裂和减数分裂中忠实地传递，并包含着丝粒和通常限于内源着丝粒存在的动粒蛋白（Blower and Karpen 2001）。然而，新着丝粒形成需要接近内源着丝粒（420kb），如倒位衍生出的 γ238；而且，当着丝粒周边异染色质出现时，新着丝粒不能在原着丝粒的任何一边形成（Maggert and Karpen 2001）。这些结果表明，新着丝粒通过原着丝粒染色质表观遗传延伸到邻近常染色质形成，随之发生着丝粒特征和功能性的表观遗传延伸。这一过程受异染色质阻滞并与 CENP-A 过表达异位合并入常染色质（而非异染色质）是一致的（Heun et al. 2006）。并且说明了着丝粒染色质的范围是由两个表观遗传过程决定的：CENP-A 填充和延伸以及异染色质形成或阻抑。

过表达的研究，CENP-A 不能定位到非着丝粒区域产生了功能异常的动粒（Heun et al. 2006）。因为这一组蛋白变体对于着丝粒功能是必要的，我们特别定义 CEN 染色质为 DNA 和 CENP-A 相关蛋白区域。

包含 CENP-A 的核小体在结构上与含有常规 H3、H2A、H2B 和 H4 的组蛋白核心有不同之处。CENP-A 核小体能够在体外由纯化的 CENP-A 和组蛋白 H2A、H2B 和

H4 构成，与以前观察的结果一致，它们在体内是同型的，例如，它们含有两个 CENP-A、一个以上的 H3 和一个 CENP-A（Yoda et al. 2000）。详细的生物物理分析表明 CENP-A 和 H4 的界面与 H3-H4 的界面不同，在具有内源 CENP-A 的情况下，H4 的相互作用域足以把 CENP-A 定位到着丝粒（Black et al. 2004）。

最初，着丝粒核小体 H3 被 CENP-A 取代，表明 CENP-A 组成了与动粒相关的所有染色质。在裂殖酵母中，CENP-A 贯穿 5～7kb 的核心区域均匀分布。包含 H3K9 甲基化和异染色质蛋白的大异染色质域分布在这些核心的侧位（Pidoux and Allshire 2004）。然而，详尽的细胞学和免疫沉淀反应研究揭示着丝粒染色质在多细胞真核生物中有更复杂的组成和组织。果蝇、人类和鼠着丝粒散在的 H3 的区块，含有核小体的 CENP-A（统一称为 CEN 染色质）侧位有甚至更大（数百 kb 到 Mb）的着丝粒周边异染色质区块（图 14-6b）（Blower et al. 2002；Yan et al. 2005）。

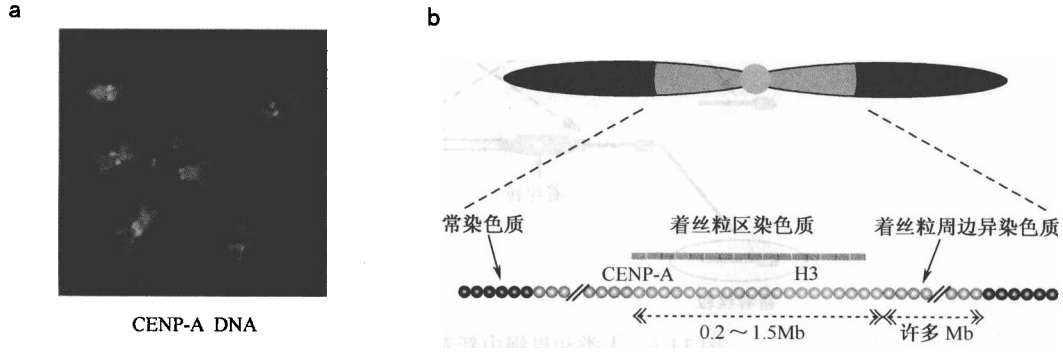



图 14-6 着丝粒染色质的组织

(a) CENP-A 是一个高度保守的着丝粒特异的组蛋白变体。图中显示了果蝇有丝分裂染色体中着丝粒的定位。(b) 果蝇和人类的 CEN 染色质包含间隔分布的 H3 和 CENP-A 核小体区域蛋白，其两侧具有着丝粒周边异染色质（Blower et al. 2002）。

H3 和 CENP-A 域的散布引出了一个关于 CEN 染色质的表观遗传特性的关键问题。特别是，多细胞真核生物的 CEN 染色质 H3 子域的修饰类似常染色质还是异染色质？或者它们有自己独特的标记？而且，较大的真核着丝粒每个散在的 CENP-A/ H3 单位与单个裂殖酵母着丝粒相当吗？这些问题在验证散在的 H3 和 CENP-A 核小体区块的特异转录后修饰时得以强化，因为这揭示了更大的复杂性。令人吃惊的是人类和果蝇中散在的 H3 域含有一个通常与常染色质相关的标记物 H3K4me2，而缺少侧位异染色质相关的 H3K9me2 和 me3（图 14-7a）（Sullivan and Karpen 2004；Lam et al. 2006）。然而，像异染色质，散在的 H3 核小体中，多个形式的 H3 和 H4 乙酰化都是缺失的，这与 H3K4me3 一样。因此，CEN 染色质内的 H3 核小体起到了一种有别于规范的常染色质或异染色质的修饰模式（图 14-7b）。这些结果也表明果蝇和人类的着丝粒不是由简单的重复的裂殖酵母样着丝粒构成的。然而，着丝粒区域的整体组织的保守性是很重要的，因此完整的 CENP-A 染色质区域侧位是着丝粒周边异染色质，在所有的多细胞真核生物和裂殖酵母中它们都有 H3K9 甲基化。

组蛋白修饰在 CEN 和侧位染色质中起到的是什么样的功能？在 CEN 区域不同的

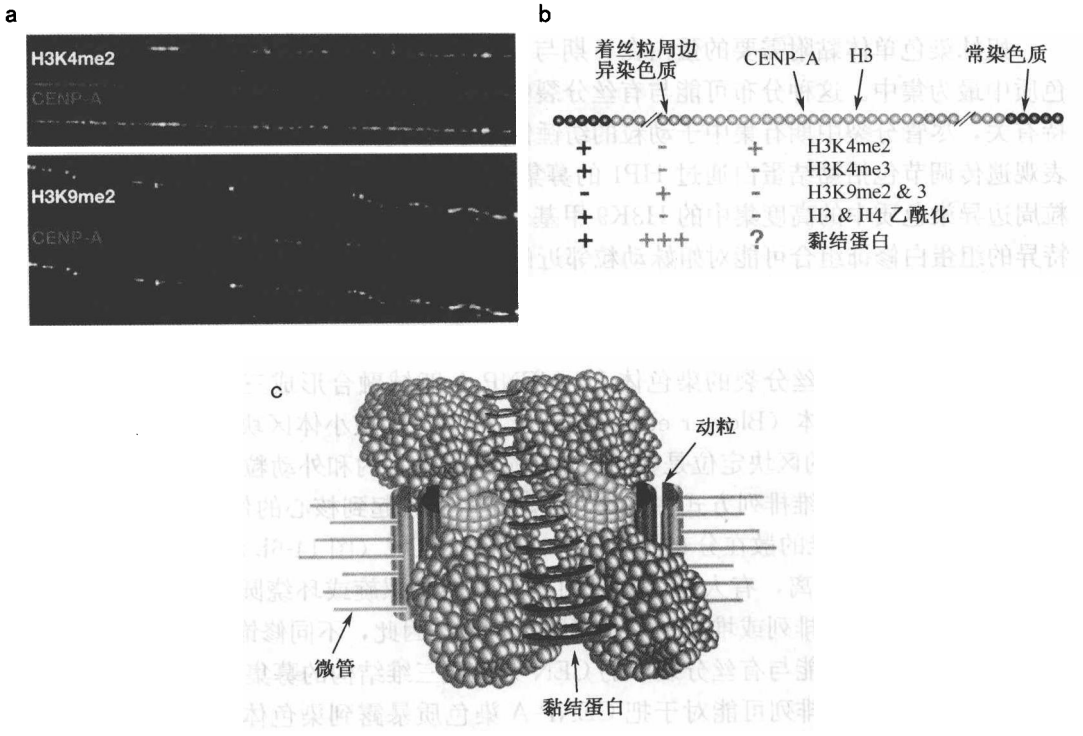



图 14-7　着丝粒染色质的独特的组蛋白修饰方式

(a) 采用免疫荧光实验应用特异的识别延伸的染色体纤维上组蛋白修饰的抗体进行实验，结果表明散在的含有 H3 的核小体区块具有一种不同寻常的常染色质和异染色质的修饰方式 (Sullivan 和 Karpen 2004)。例如，尽管着丝粒在大多数真核生物中都是嵌入大的着丝粒周边异染色质中，但含有 H3K9me2 修饰的散在的 H3 区块正常情况下却与“开放”常染色质（顶部）相关联，并且着丝粒周边侧位区域（底部）缺少异染色质标志 H3K9me2。(b) 以在果蝇和人类中延伸的染色质纤维研究为基础的着丝粒染色质二维结构示意图。＋和－分别表示为组蛋白修饰在常染色质和异染色质中的有和无以及 H3 核小体在着丝粒染色质上散在的区块 (Sullivan and Karpen 2004; Lam et al. 2006)。(c) 有丝分裂染色体着丝粒区域染色质三维组织结构模型。相似修饰的核小体之间的关联可能有助于着丝粒和侧位染色质的不同的三维结构的形成。散在的 CENP-A/CID 和不同修饰的 H3 和 H4 可能介导了中期染色体中发现的圆柱形三维结构的形成 (Blower et al. 2002; Sullivan and Karpen 2004)。H3K9me2 染色质，其能够募集异染色质蛋白（如 HP1）和黏附蛋白（如 RAD21/SCC1），出现在位于有丝分裂姐妹染色单体之间的内动粒和侧位着丝粒染色质区域。这种排列方式可能对于 CENP-A 朝向有丝分裂染色体的极向面以及便于外部动粒蛋白的募集、促进 HP1 自身相互作用以及正确的染色体浓缩和黏附是必需的。黏结蛋白呈现为环形结构与最近提出的模型一致。

染色质状态可能与着丝粒区域的特性有关，如 CEN 和侧位异染色质的不同复制时间 (Sullivan and Karpen 2001; Blower et al. 2002)。着丝粒周边侧位异染色质修饰可能也通过制造 CEN 染色质的障碍来维持着丝粒的大小。在果蝇中，当侧位异染色质移除之后 CEN 染色质稳定地扩展到邻近的序列，使新着丝粒具有活性（图 14-5b）(Maggert and Karpen 2001)，CENP-A 的过表达导致在常染色质中异位产生，但是异染色质中则正常 (Heun et al. 2006)。有趣的是，CENP-A 在人类细胞中的过表达导致 CEN 染色质的蔓延和侧位 H3K9 甲基化的改变 (Lam et al. 2006)。因此，着丝粒的大小可能是

由两个表观遗传状态的平衡决定的：CEN 染色质和侧位着丝粒周边异染色质。

姐妹染色单体黏附需要的蛋白在 S 期与 DNA 复制共同完成，在着丝粒周边的异染色质中最为集中。这种分布可能与有丝分裂中动粒的双向定位及分裂中期黏附状态的维持有关，尽管分裂中期有集中于动粒的纺锤体介导的力存在（Watanabe 2005）。黏附的表观遗传调节包括黏结蛋白通过 HP1 的募集（裂殖酵母的 Swi6），这反过来又被着丝粒周边异染色质中的高度集中的 H3K9 甲基化介导（Nonaka et al. 2002）。因此，CEN 特异的组蛋白修饰组合可能对姐妹动粒邻近的异染色质黏附复合物的募集有重要作用，从而确保黏附区域和动粒区域在空间上的分离（Sullivan and Karpen 2004；也见第 6 节）。

在人类和果蝇有丝分裂的染色体上，CENP-A 亚域融合形成三维的圆柱形结构去除了大部分 H3 核小体（Blower et al. 2002）。CENP-A 核小体区块是面向染色体的极向面的，H3 核小体的区块定位是朝向内染色质区域的。内和外动粒蛋白是包绕 CENP-A 圆柱体的；这一三维排列方式在募集其他动粒蛋白时起到核心的作用（Blower et al. 2002）。为了顺应二维的散在分布的 CENP-A 和 H3 区域（图 14-6b 和 14-7b）与三维的有丝分裂染色体的分离，有人认为 CEN DNA 可能是螺旋或环绕圆柱状结构的，导致与相同的成分呈线性排列或堆积在一起（图 14-7c）。因此，不同修饰的散在的 H3 核小体和侧位异染色质可能与有丝分裂中的 CEN 染色质三维结构的募集有关（Sullivan and Karpen 2004）。这种排列可能对于把 CENP-A 染色质暴露到染色体外部有关，它能通过建立纺锤体极相关的双向姐妹动粒方式募集动粒蛋白。

### 3.4 着丝粒结构，功能和传递模型

现在正在探索的关键问题是 CENP-A 和其他表观遗传标记是如何特异的在着丝粒定位和传递的？另一种思考这个问题的方式是分析 CENP-A 为什么能够仅在着丝粒染色质进行装配？一个很有吸引力的模型提出 CEN DNA 复制和 CENP-A 表达与大块染色质调节的 CENP-A 在着丝粒的特异整合，二者有不同的时间安排（Ahmad and Henikoff 2001）。然而，CEN DNA 的复制贯穿于人类和果蝇 S 期，与大多数 DNA 复制同时发生（Shelby et al. 2000；Sullivan and Karpen 2001），CENP-A 的整合发生在非 DNA 复制阶段（Shelby et al. 2000；Ahmad and Henikoff 2001）。这些发现排除了 CENP-A 传递和着丝粒特征性的严格复制时间机制。

一个有趣的发现是 CENP-A 通过一个复制非依赖的方式经组蛋白交换复合体主动整合进核小体（Shelby et al. 2000；Ahmad and Henikoff 2001），由此提出了一个更具吸引力的机制。H3.3 是 H3 的一个变体，它的复制非依赖组装（Ahmad and Henikoff 2002）是由一个称为组蛋白调节因子 A(histone regulator A，HIRA）的复合体介导的，而非由常规 H3 核小体复制依赖的染色质组装因子（chromatin assembly factor，CAF）复合体介导（Nakatani et al. 2004）。HIRA 成分的缺失导致酿酒酵母中 CENP-A 的错误定位（Sharp et al. 2002）。然而，在多细胞生物和裂殖酵母中HIRA的组分是否影响着丝粒染色质还不清楚，在这两者中着丝粒特征性是表观遗传决定的。更重要的是，这些蛋白在染色质组装和结构总体上起一定的作用，如 H3.3 的置入；因此，已经明确的

HIRA 成分不能解释其在着丝粒的特异整合。有一种可能是 HIRA 复合体的一小部分，包含着仅与 HIRA 作用的因子，其识别复制的 CEN 染色质中已经存在的 CENP-A 核糖体；但是，目前还没有鉴别出这样的特异因子。一种非特异组装因子参与调节的设想是由 CENP-A 或 CENP-A 核小体提供了特异性。例如，已明确的 CENP-A 和 H4 在结构上的关系能够为新 CENP-A 在着丝粒的装配提供特异性，正如 CENP-A 定位到着丝粒的相互作用域的现象所提示的那样。(Black et al. 2004)。然而，还不清楚这些结构域在内源 CENP-A 缺乏时是否足够定位到着丝粒。

目前需要从新的角度思考有关着丝粒特征和传递的表观遗传的调节的问题。裂殖酵母中，CENP-A 蛋白降解缺陷导致其错误整合入正常的非着丝粒区域，正常情况下这些除去内源着丝粒的区域应该是被一种未知的“清除”机制所移除（Collins et al. 2004)。这提示着丝粒特征性的调节可能是在核小体装配后的某一时间进行的。然而，CENP-A 在果蝇中的错误定位导致异常的动粒形成（Heun et al. 2006)，表明移除错误定位的 CENP-A 可能是裂殖酵母特异的。

然而，这种“负特异”模型的其他形式也是值得思考的。对所有着丝粒特征性模型的关键问题是什么提供了这种特异性？在这一事件中，为什么 CENP-A 等蛋白只保留在了一个位点？由这一事实提出的一个新的思路是其与纺锤体稳定的联系可能是功能着丝粒/动粒特有的性质（Mellone and Allshire 2003)。因此，着丝粒延伸和 CENP-A 整合位点可能是在有丝分裂过程中决定的，机制是一种对动粒-纺锤体接触或纺锤体介导的张力的感知。另一个值得探讨的思路是散在的 H3 核小体经组蛋白修饰酶（如乙酰转移酶、甲基转移酶和激酶）修饰可能有助于着丝粒特征性的蔓延，其与 CENP-A 相关蛋白结合或代替 CENP-A 相关蛋白（Sullivan and Karpen 2004)。不同修饰的散在的 H3 亚域（图 14-7）可以产生新 CENP-A 装配必需的“允许”状态的染色质结构。

不考虑模型的因素而鉴别 CENP-A 置入着丝粒所需因子有可能提供新的视角。生物化学和遗传学研究鉴定出一些影响 CENP-A 在着丝粒信号的因子，包括已知的在非复制依赖染色质装配中涉及的因子。然而，没有任何一个因子能够特异的与 CENP-A 或其他着丝粒染色质蛋白或修饰相互作用。尽管如此，这些因子被鉴定就是令人兴奋的成果，阐明特异的机制应该也不远了。

### 3.5　表观遗传学与着丝粒的演变

着丝粒对于细胞和生物体生存的重要性是不言而喻的，着丝粒应该不会出现获得或丢失的问题。如果细胞或单细胞生物体的着丝粒“硬线”（hard-wired）地存在于 DNA 的一级序列中具有显著优势，那么为什么着丝粒要受到表观遗传机制的调控？

一个非常值得探讨的问题是着丝粒特征性的表观遗传调控对进化中累积于染色体、DNA 序列和蛋白质的变化是必要的。哺乳动物（如灵长类和有袋类动物)、昆虫和其他门类的研究表明，着丝粒的获得或缺失是染色体进化的标记（Ferreri et al. 2005)。相近物种染色体臂在排列和关联上经常是不同的，即使是在 DNA 序列基本一致的基础上。着丝粒的获得或缺失经常伴有转位和重排，但也可能是受到他们的影响。例如，对仅存在一个着丝粒的需求会引起许多无意义的双着丝粒和无着丝粒染色体重排，除非发

生着丝粒失活或新着丝粒活化（图 14-8a）。因此，将着丝粒通过蔓延、跳跃或活化，从一个 DNA 序列转移到另一个位置的能力，可能会加速染色体的进化。另外，扩张、收缩和碱基改变在进化过程的高度重复卫星序列中经常发生，并且与着丝粒部位有一定联系（图 14-8b）。尤其是植物的着丝粒在进化过程中表现出快速和显著的 DNA 序列和位置的变化（Hall et al. 2006）。然而，现在还不清楚是否是 DNA 的变化引起着丝粒的移动，还是对着丝粒运动的反应，或是根本不相关。尽管如此，仍然需要有表观遗传提供的不依赖于序列改变的维持着丝粒功能和蔓延的机制。

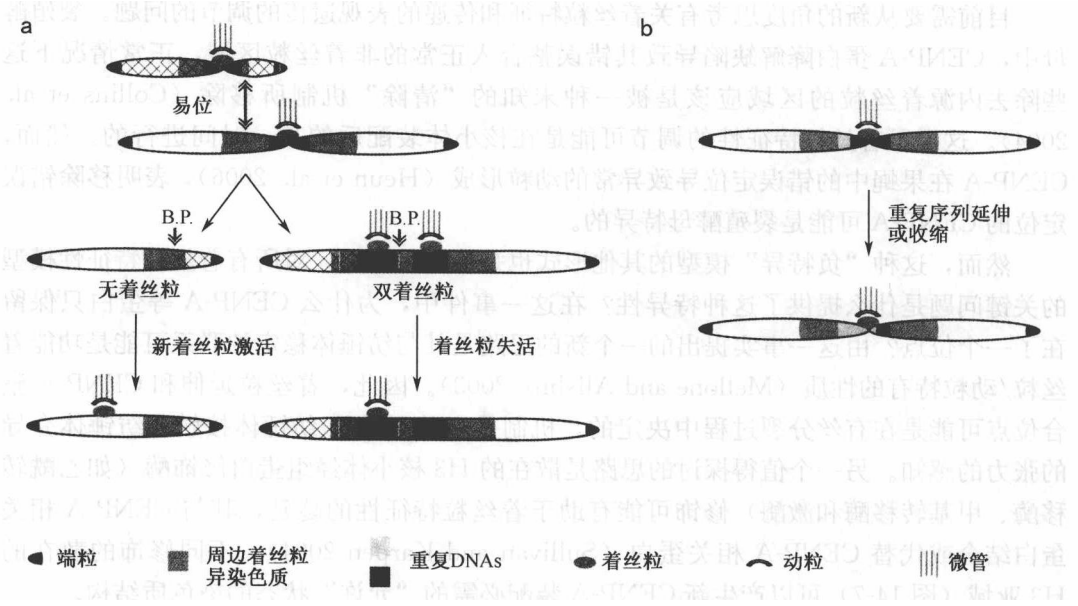



图 14-8 染色体进化和着丝粒特征性表观遗传调控

与特异 DNA 序列相关的着丝粒可塑性可能对于进化中调节序列改变和染色体重排是必要的。（a）进化中经常发现的易位能够产生无着丝粒和双着丝粒染色体，两者正常情况下在有丝分裂和减数分裂过程中会丢失。表观遗传调控和可塑性使得无着丝粒片段获得新着丝粒功能，双着丝粒染色体上的一个着丝粒失活，或产生两个易位产物的正常遗传。（B. P. 为易位分界点）（b）大多数真核生物的着丝粒镶嵌在异染色质和重复 DNA 序列之中，特别是高度重复的卫星序列。这些序列在进化中发生变化的几率非常高，经常发生碱基改变或排列扩张或收缩。着丝粒对特异 DNA 序列严格的依赖性会导致着丝粒和动粒功能的丢失和有害的染色体丢失。相反，着丝粒特征性和位置的表观遗传调控提供了一种尽管有序列改变也能维持着丝粒的机制。

基于跨物种着丝粒蛋白的氨基酸序列比对提出了一种新的说法，即结构域与 CEN DNA 结合“共进化”以适应卫星序列的变化（Malik and Henikoff 2002；另见第 13 章）。这一模型的一个有趣的成分提示这些改变是由减数分裂驱动的，是一个减数分裂过程发生的等位基因的选择（并且是缺乏整体染色体时）。因此，具有对于 CEN 染色质蛋白最强亲和力的一个着丝粒，在整合入功能性生殖细胞这一点来说是最成功的（如在雌性减数分裂中是卵母细胞而不是极体）（见第 5 节中有关减数分裂驱动的其他例子）。这样反过来又能促使其他着丝粒适应相同的 DNA 序列和蛋白变体从而能够有效地分离。

表观遗传标记，如 CENP-A 的缺失，提供了另外一种在着丝粒 DNA 不去除的情况

下着丝粒失活的机制，如同大量生物中 CENP-A 缺失的结果所描述的那样（Cleveland et al. 2003）。着丝粒获得机制的探讨更具有挑战性，这是因为其要求在 DNA 序列不变的情况下寻找表观遗传标志。果蝇中，试验性诱导新着丝粒的研究指出了一个着丝粒获得的分子机制，即通过 CENP-A 等关键着丝粒蛋白的顺式蔓延（图 14-5b）（Maggert and Karpen 2001）。然而，这一模型并不适用于人类新着丝粒的形成和进化过程中获得着丝粒的多数情况，因为它们都是在亲本着丝粒较远处发生的。对于这些着丝粒生成更适合的机制包括由于过表达导致的 CENP-A 误定位，在果蝇试验中有一样的结果即易位动粒的形成（Heun et al. 2006）。CENP-A 过表达在结肠癌和乳腺肿瘤中都有所发现，表明可能这与癌症中整体基因组不稳定有关。进一步对 CENP-A 误定位在不同类型人类肿瘤中的发生率以及在癌症起始和进展过程中出现的时间选择研究能够直接验证这一假说以及其在着丝粒进化中的作用。

最后，散漫着丝粒染色体可能可以代表第一批着丝粒；动粒形成可能首先随着随机序列特异性进化，随后由于转座子侵入，卫星 DNA 的扩展以及侧位异染色质的形成，进化为单着丝粒染色体。然而，全着丝粒染色体也可能是从单着丝粒染色体进化而来，机制是异染色质边界元件的丢失和着丝粒染色质的顺式蔓延，这个过程与果蝇新着丝粒产生相类似（Maggert and Karpen 2001）。

### 4. 异染色质和减数分裂配对/分离

大多数生物中减数分裂由以下步骤构成：①配对，同源染色体排列；②联会，同源体通过联会复合体（synaptonemal complex，SC）连接；③重组，从空间结构上连接同源染色体并进行遗传信息交换；④减数分裂Ⅰ期同源染色体向相对的两极分离；⑤减数分裂Ⅱ期姐妹染色单体分离（图 14-9）。异染色质元件如着丝粒和端粒在控制常染色质重组发生中起到重要作用；可以肯定的是，异染色质成分（特别是着丝粒）对于分离至关重要。而且，许多最近的研究论证组蛋白修饰或其他的表观遗传成分和过程在促进减数分裂过程中起到重要作用。例如，秀丽隐杆线虫中 H3K9 甲基化需要的 HIM-17 蛋白，对于起始减数分裂重组的 DSB 的形成是必需的（Reddy and Villeneuve 2004）。组蛋白激酶 1（histone kinase，Hsk1）对于裂殖酵母中 DSB 的起始也是同样需要的。组蛋白甲基化对于果蝇（Cullen et al. 2005；Ivanovska et al. 2005）和哺乳动物系统（De La Fuente 2006）中减数分裂染色体分离过程也是需要的。

虽然利用重组以确保减数分裂的染色体分离是基本通用的，减数分裂系统在缺乏重组的同源体中也能存在（即无交叉的减数分裂系统）。在这样的系统中，异染色质配对以及其他表观遗传修饰的底物，似乎都起到关键作用。这在果蝇中尤为明显，在酵母中可能也是事实。最后，许多减数分裂驱动系统已经被了解清楚——如减数分裂分离过程对一个等位基因的偏爱，其中，特定的异染色质元件的修饰可引起或允许后续特异染色体或整套染色体的丢失或失活。

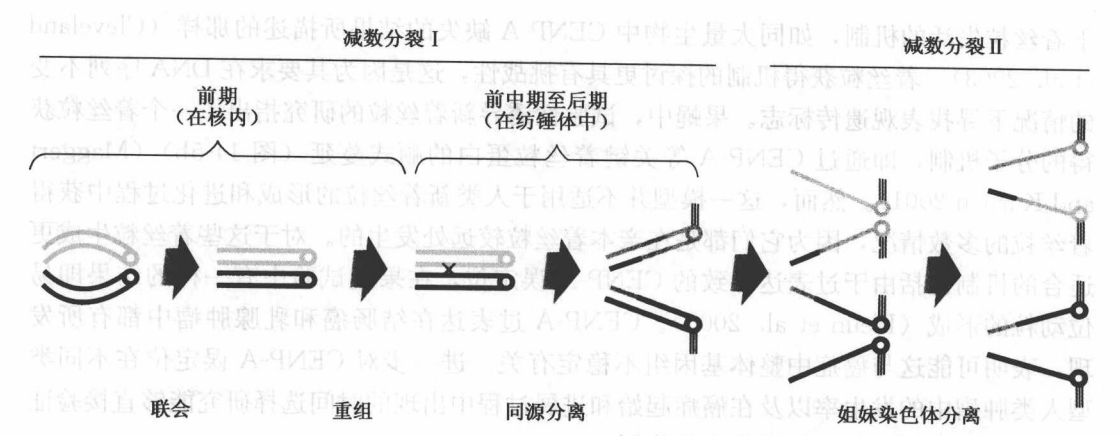



图 14-9 从力学角度阐述减数分裂过程

在第一次减数分裂过程中，同源染色体配对必须完成三件事情。首先，同源染色体必须沿着总体长度进行配对。实际上，对于所有的生物体，配对以一种紧密的状态达到完善，同源染色体相互连接。这种状态称为联会。第二，同源染色体通过重组彼此锁在一起，称为交换（crossing-over）。这些结构上的交换（crossover）称为交叉（chiasmata），即在交叉两边的姐妹染色单体上出现结构上的互锁。配对和重组都出现在前期（在核膜破坏之前）。第三个主要事件是分离，是在核膜破裂之后产生的。在纺锤体装配的早期（前中期），染色体聚集产生赤道板。在许多雄性动物中，具有中心粒减数分裂纺锤体，在多数雌性动物中，纺锤体则无中心体。在这种情况下，染色体本身将形成一个团块最终变成赤道板，并环绕组成一个双极染色体。一旦染色体确定了正确方向（即同源动粒达到纺锤体两极并在赤道板得以平衡），各种机制触发了后期的开始。后期，姐妹染色单体黏附沿着染色体臂释放（但不靠近着丝粒处）。这时联会解体，交叉（chiasmata）由交换（crossover）产生，导致同源染色体分离并在后期 I 向两极移动。减数分裂 II 期，基本上就等于没有复制和重组的单倍体有丝分裂。

## 4.1 雄性果蝇的一个异染色质配对位点的发现

第一个介导同源体配对和分离的异染色质元件是介导雄性果蝇精子发生过程中无交叉性染色体对配对和分离的着丝粒周边异染色质结构。该结构称为配对区（collochores），是与 rDNA 重复序列相对应的；这些 rDNA 基因整合入配对区缺失的 X 染色体后，在 rDNA 插入位置可恢复 X-Y 配对（McKee 1998）。与配对相关的至关重要的 rDNA 重复片段是基因间隔内的 240bp 重复序列。当在多个副本中出现的时候，这个 240bp 重复序列有助于雄性果蝇减数分裂的 X-Y 染色体配对以及后续的分离。虽然果蝇精母细胞中异染色质联结促进常染色质配对是仍待解决的问题，我们将在下文中讨论大量的介导或维持配对缺失所引起常染色质不正确的激活或失活的发现。这种情况经常会导致精子发生的失败（McKee 1998）。

## 4.2 雌性果蝇异染色质配对促进分离

在雌性果蝇中有两种主要机制确保同源染色体分离：交叉系统介导交叉同源体分离，该过程发生交换；无交叉系统可保证配对同源体分离但无法发生交换。例如，小点状四号染色体从不发生交换，而且它们在减数分裂时彼此分离并具有高保真性。细胞学和遗传学试验证实无交叉配对和分离需要异染色质。无交叉染色体同源异染色质区域配对从前早期一直持续到前中期 I，无交叉同源体开始从它们的配对体分离（图 14-10）

(Dernburg et al. 1996)。另外，缺失研究发现雌性果蝇同源无交叉分离是完全异染色质同源性依赖的（Hawley et al. 1993；Karpen et al. 1996）。介导这些配对和分离是否是异染色质序列或异染色质蛋白的一般特性，或是散在于异染色质内的众多特异序列的结果，还尚不清楚。

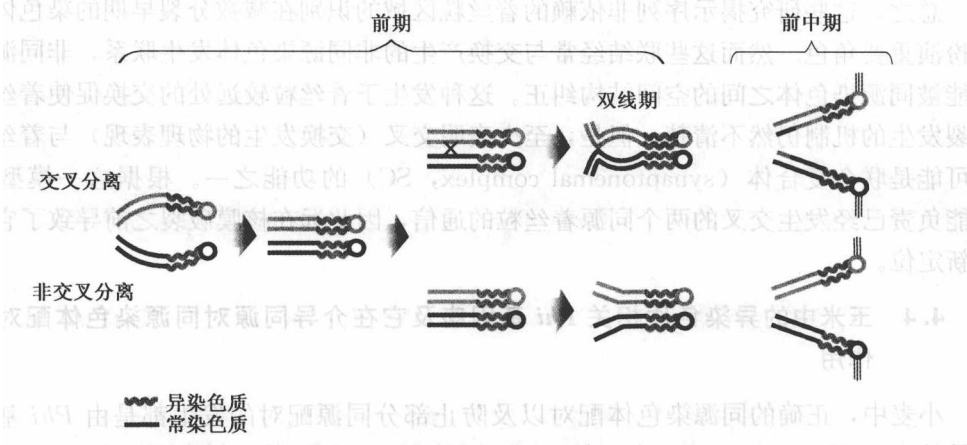



图 14-10　雌性果蝇中确保同源体分离的两种机制

雌性果蝇的减数分裂过程具有一个非正常的双线期（下）。在规范的减数分裂中，双线期（有时候也称为双线期-终变期）为前期的最后一个阶段，其间同源体互相排斥仅有交叉互相连接（上）。然而，在雌性果蝇，只有染色体的常染色质臂在前期末排斥或分离；异染色质区域在前期末（核膜裂解）和前中期也紧密结合配对（Dernburg et al. 1996）。在无交换时这些异染色质配对对于确保分离忠实性是必要且充分的（Hawley et al. 1993；Karpen et al. 1996）。

## 4.3　着丝粒在芽殖酵母无交叉分离中的促进作用

虽然雌性果蝇侧位异染色质区域看似对于介导无交叉的分离是足够的，但是在酿酒酵母中，着丝粒序列本身即可以完成这一功能。Dean Dawson 和合作者在带有一对部分同源的染色体的酵母中分析减数分裂，这些染色体来自于不同种的酵母（Kemp et al. 2004）。尽管这些染色体带有相似的基因，它们的不同之处也足以避免出现交叉；而且也能够忠实地相互分离。虽然这些部分同源的染色体臂在减数分裂前期中相互联系的发生率也并不比异种染色体高，但是这些部分同源染色体具有高水平的着丝粒配对(Kemp et al. 2004)。而且，第三个非交换染色体的出现能够减少这些配对的发生，即使它带有一个不同的着丝粒区域，表明着丝粒配对是序列非依赖性的。

如果着丝粒配对不需要序列的同源性，那么为什么部分同源的染色体分离不是随机发生的呢？Kemp 等（2004）提出“是交换和交换后发生的联会，使同源着丝粒并列并阻止了着丝粒的随机配对”。依照这一模型，着丝粒可能是以一种完全的序列非依赖的方式起始配对。同源染色体臂发生交换时，它们的着丝粒从潜在的非同源的染色体联结中分离出来并强行进入到同源配对中。通过交换依赖性，着丝粒成为一种“单线态”，然后便可自由地与其他未配伍的着丝粒发生再联结，并最终与未发生交换的染色体的着丝粒配对。Kemp 等提出一个相似的过程，就像高校舞会中较有魅力的孩子找舞伴配对，然后较无吸引力的孩子作为舞伴配对。几十年以前 Rhoda Grell 就已经提出了一个

对该过程更为正式的描述。她提出一旦交换配对发生，这些可用的染色体就会自由地发生再联结（Grell 1976）。事实上，最近的研究表明，在减数分裂过程中，所有酵母染色体都是非同源联结的并且经历一个重组后再分类的过程使同源染色体联结（Tsubouchi and Roeder 2005）。

总之，这些研究揭示序列非依赖的着丝粒区域的识别在减数分裂早期的染色体联结中扮演重要角色。然而这些联结经常与交换产生的非同源染色体发生联系，非同源性联结能被同源染色体之间的空间结构纠正。这种发生于着丝粒较远处的交换促使着丝粒重分裂发生的机制仍然不清楚。但是，至少表明交叉（交换发生的物理表现）与着丝粒相关可能是联会复合体（synaptonemal complex，SC）的功能之一。根据这一模型，SC可能负责已经发生交叉的两个同源着丝粒的通信，因此远在核膜破裂之前导致了它们的重新定位。

### 4.4 玉米中的异染色质相关 *Phi* 基因座及它在介导同源对同源染色体配对中的作用

小麦中，正确的同源染色体配对以及防止部分同源配对的发生都是由 *Phi* 基因座介导的（Griffiths et al. 2006）。在 *Phi* 基因缺失时，同源染色体的配对和交换变得频繁（Luo et al. 1996）并且降低了减数分裂和有丝分裂着丝粒配对的保真性。着丝粒错误配对在更加混乱的部分同源染色体配对和交换方面的意义仍然有争议（Dvorak and Lukaszewski 2000）。但是 *Phi* 剔除的效应能够被基因组中加入异染色质额外的 B 染色体所抑制（Bennett et al. 1974）。最近，*Phi* 基因座被定位于一个 2.5 Mb 的区间中，该区间含有插入了 *cdc2* 相关基因簇的亚端粒异染色质。这一现象发生在以现代小麦进化为特征的多倍体化之后（Griffiths et al. 2006）。虽然这些关于 *Phi* 基因座的研究表明小麦减数分裂中配对选择的表观遗传调控，在介导防止部分同源配对中异染色质的作用以及同源配对的促进的机制都还不清楚。

## 5. 异染色质和减数分裂驱动

在众多例子中我们发现异染色质区域或元件在一个称为“减数分裂驱动”（meiotic drive）的过程中发挥作用。减数分裂驱动是指一个同源体以其配对体为代价加强自身传递可能性的能力，如在一个 A/a 杂合体，A 型接合体产生和利用比 a 型接合体更频繁。我们的注意力主要集中在异染色质能够引起它们的同源体丢失或破坏的例子中。虽然我们仅详细描述了黑腹果蝇中的分离紊乱因子（*segregation distorter*，*SD*）系统和蕈蚊及丽蝇金小蜂中的染色体丢失，其他的异染色质相关的减数分裂驱动的系统也存在，如雀类异染色质化和生殖细胞系限制性染色体丢失以及玉米中 B 染色体的堆积。

### 5.1 雄性果蝇分离紊乱因子

异染色质介导减数分裂驱动的令人印象深刻的例子之一是在独立于自然种群的黑腹果蝇中的一个单模标本，称为分离紊乱因子或 *SD*（Sandler and Hiraizumi 1959）。一个

带有单倍型的果蝇二号染色体在精子细胞成熟过程中能够通过阻止同源体正确的浓缩来减少它的野生型同源体（表示为 $SD^+$）。只有带有 *SD* 的精子细胞能够完全成熟为成熟的精子并且能够传递给后代。

*SD* 染色体带有几个分散的与其功能有关的遗传元件。其中最重要的就是常染色质 *SD* 突变本身，它通过作用于一个携带分开的异染色质目标元件的 *Rsp* 序列引起精子细胞功能障碍（Ganetzky 1977）。*Rsp* 自身由一个位于的二号染色体的着丝粒周边异染色质的重复元件构成（Wu et al. 1988）。同源的 $SD^+$ 染色体浓缩缺陷的发生率依赖于 *Rsp* 重复序列拷贝数的多少。除此之外，还有许多其他的元件有利于 *SD* 发挥引起紊乱能力，通常位于携带 *SD* 的染色体之上。

*SD* 突变本身是一个包含 *RanGAP* 基因的小串联复制。重排产生了一个羧基端缺失了 234 个氨基酸残基的 RanGAP 蛋白突变体，从而导致核转运缺陷（Kusano et al. 2003）。为什么这样一个突变可能引起带有 *Rsp* 的染色体产生染色质浓缩失败仍然不清楚（见 Kusano et al. 2003 中讨论的可能机制）。

## 5.2　蕈蚊中父源染色体丢失和响应元件定位

蕈蚊的体细胞和生殖细胞内均发生父源染色体剔除的复杂过程。它们都来源于 $A_mA_pX_mX_pX_p$ 基因型的合子，即从母亲获得一套常染色体和一个 X 染色体，从父亲获得一套常染色体和两个 X 染色体。注定成为雌性的胚胎体细胞将丢失一个来自于父亲的 X 染色体，注定成为雄性的胚胎体细胞将丢失两个来自于父亲的 X 染色体。关于 X 染色体剔除的实际机制，涉及不完全姐妹染色单体分离的理论，由 de Saint Phalle 和 Sullivan 在 1996 年提出。

父源 X 染色体丢失也发生在由生殖细胞系组成的细胞中。在生殖细胞系发展的后期阶段，一个单独的 $X_p$ 染色体在两性中都将被遗弃。父源 X 染色体丢失后，整个父源染色体分布表现为去浓缩化，实验时染色也就很浅。浓缩的不同将会维持直到两性一龄幼虫后期，也就是下颌角有丝分裂的开始。在这一阶段，父源和母源的染色体都表现为完全的浓缩。随后发生的雌性减数分裂具有正常的细胞学特征，而雄性减数分裂从两个方面来说是不正常的。首先，整套的父源染色体是被剔除的；其次，在减数分裂Ⅱ期母源 X 染色体不分离从而形成一个具有一套常染色体和两个 X 染色体的精子细胞。

有趣的是，组蛋白 H3 和 H4 乙酰化与这些形式的染色体剔除有关（Goday and Ruiz 2002）。首先，在生殖细胞系发育的早期，只有一半的来自父源性的染色体在 H3 和 H4 被高度的乙酰化。正如 Goday 和 Ruiz 在 2002 年提出的，“生殖细胞系染色体不同的乙酰化标签与父源和母源早期生殖细胞系基因组染色质的数目不同相一致。”这一原则的例外是在生殖细胞发展过程中，父源的 X 染色体将被剔除，但发生高度的乙酰化。在下颌角有丝分裂开始之前，所有父源和母源的染色体均高度乙酰化。最有意思的是，在雄性的减数分裂中，整套的父源染色体都将丢失，整套的父源染色体的 H3 和 H4 乙酰化水平较低，这是较相对应的母源染色体高度乙酰化而言的。虽然组蛋白修饰的变化与后续的剔除的相关机制仍然不清楚，但这种不同似乎能标记有丝分裂和减数分裂的染色体剔除。

### 5.3 丽蝇金小蜂中父源染色体丢失

在营寄生生活的丽蝇金小蜂中，未受精（单倍体）卵发展成雄性，而受精（二倍体）卵通常发展成雌性（Beukeboom and Werren 1993）。然而，存在一种额外的 B 染色体，或称为 *PSR* 染色体，它能使受精卵发展成雄性。*PSR* 染色体在受精卵中通过超浓缩和丢失所有父源染色体（除了它自身）实现这种性别的转变（Werren et al. 1987）。*PSR* 染色体大部分都是异染色质并含有常染色体上没有的串联重复序列。缺失-定位这一染色体的努力确实产生了非功能性的 *PSR* 染色体并能传递到子代中（Beukeboom and Werren 1993）。然而，这些研究都没有能够鉴别出这个染色体上与之触发父源基因组破坏相关的特异区域（或重复序列）。但是，比较著名的是在一个细胞系中此类剔除衍生物可由无功能恢复到有功能。Beukeboom 和 Werren(1993) 提出 *PSR* 染色体上的某些重复单位可能对于受精后父源染色体剔除所需因子来说是一个“陷阱”。对于丽蝇金小蜂性别决定中基因组印记的研究表明，父源和母源之间确实存在“状态差异”，这可能是 *PSR* 染色体的贡献（Dobson and Tanouye 1998）。

### 5.4 玉米中的异染色质 knob 10 序列在促进减数分裂I染色体分离中的作用

另一个异染色质相关减数分裂驱动的例子在保留有异染色质插入的染色体中被发现，称为“knob”（伪着丝粒）；它们在雌性减数分裂过程中有利于带有 knob 的染色体传递。Knob 的分子组成包括两套串联重复序列，都以一种长而无间断的排列方式出现（Dawe and Hiatt 2004）。异染色质 knob 的功能作为伪着丝粒有赖于 10 号染色体的一个变体，称作异常染色体 10（Ab10）。当 Ab10 出现的时候，knob 形成伪着丝粒在内源着丝粒之前向着第一次减数分裂纺锤体极移动，并且它们仅外侧与微小管联合，不是完全功能性内源着丝粒的“端点”（end-on）联合方式。因此，尽管命名有混淆，但由于 Knob 本身不能推动正常的分离以及不能募集蛋白或表现与内源着丝粒和动粒相关的功能，它们不是“新着丝粒”（见 3.3 节）。

在减数分裂驱动方面，性母细胞中与特定 knob 杂合的交换导致一个仅在两个单体中的一个上带有 knob 的交叉产物的形成（Rhoades and Dempsey 1966）。伪着丝粒 knob 指引带有 knob 的染色单体到四个减数分裂产物之一，称为基本大孢子，是两个“外边”核之一（属于四个减数分裂产物），也是唯一可以受精的（Rhoades and Dempsey 1966；Yu et al. 1997）。因此，这一系统中的减数分裂驱动是由异染色质介导的一个姐妹染色单体定向分离到功能性接合体，这是以损害另一个姐妹染色单体为代价的。表观遗传机制如何调节 knob 行为和伪着丝粒活性目前还是未知的。

## 6. 减数分裂过程中非配对 DNA 导致的基因沉默

许多例子表明减数分裂配对失败导致未配对的 DNA 表观遗传沉默。在链孢霉的例子中，这种沉默一般限制在未配对的区域。

### 6.1 链孢霉减数分裂中非配对 DNA 导致的减数分裂沉默

一个特殊的减数分裂期间基因沉默的例子，称为未配对 DNA 导致的减数分裂沉默

(meiotic silencing by unpaired DNA，MSUD)，报道于链孢霉中（参见综述 Hynes and Todd 2003 ）。这一过程可使某基因未配对的副本沉默其本身或这一基因任何其他已配对的副本（图 14-11）。这一过程通过子囊孢子成熟 1 基因（*ascospore maturation*，*asm-1*）杂合性缺失而鉴定。然而，一个未配对的序列必须足够大并与这一基因的mRNA产物有一定的同源性才能起始沉默（Lee et al. 2004）。但是，这样的序列无需包含问题基因的起始原件。沉默限制了基因中未配对的序列，不能蔓延到邻近配对的基因部分（Kutil et al. 2003）。*sad-1* 基因编码的蛋白与 RNAi 通路中的 RNA-依赖的 RNA 聚合酶（RNA-dependent RNA polymerases，RdRP）具有较大的同源性，其突变能够阻抑 MSUD 的过程（Shiu and Metzenberg 2002），这表明“MSUD 需要双链 RNA 的合成和可能的 RNA 扩增”（Hynes and Todd 2003）。虽然在减数分裂中正常表达的基因中发现过 MSUD，但在仅表达于无性生殖细胞的基因中却没有发现，更有力的提示了 RNAi 的参与（Hynes and Todd 2003）。对此说法的进一步支持源于 MSUD 过程中对类 Argonaute 与类 Dicer 蛋白的需求的发现（Lee et al. 2003）。

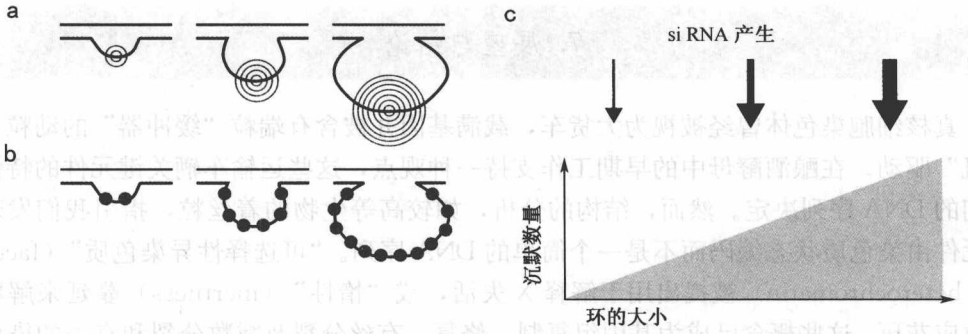



图 14-11　未配对 DNA 转录如何导致减数分裂中不同程度沉默的模型

(a) 一个单独的转录复合物在较大的环中比较小的环中有更强的转录活性，如图中同心圆所显示的。(b) 较大环可能比较小环含有更多的转录复合物（小球）。(c) 在两个模型中未配对的 DNA 在大环中将产生更高浓度的 siRNA（以箭头的粗细表示），并导致更多的沉默（经允许选自 Lee et al. 2004）。

可以认为 MSUD 是这两部分过程的交点：反式感知和 MSUD(Pratt et al. 2004)。反式感知似乎是 DNA 甲基化依赖性的，它允许一个既定的 DNA 序列“感知”减数分裂前期环境中配对区域的特性或同源性。一旦未配对序列通过这一方式被识别，MSUD 过程就会启动并抑制这些序列的表达。

未配对 DNA 减数分裂中沉默的一个 RNA 依赖机制在秀丽隐杆线虫中也被发现。DNA 在减数分裂中缺乏配对对象则成为组蛋白 H3K9 甲基化的靶点（Bean et al. 2004）。这一沉默机制需要 RNA 依赖的 RNA 聚合酶（EGO-1），但不需要剪切酶 Dicer (Maine et al. 2005)

## 6.2　小鼠中非联会染色体沉默

两种性别的小鼠中都发现了未配对 DNA 的减数分裂沉默（Baarends et al. 2005；Turner et al. 2005）。特别是雄性减数分裂中未配对的常染色体区域，或雌性减数分裂

中未配对的 X 染色体，它们都是转录水平沉默或具有组蛋白 H2A 泛素化标记，也都有雄性减数分裂细胞出现的正常 XY 体的特征（Baarends et al. 2005）。也有报道来自鼠转位杂合子的未配对 DNA 被 BRCA1 募集的检验点激酶 ATR 沉默，此过程类似于正常雄性减数分裂的 X 和 Y 染色体沉默（Turner et al. 2005）。

### 6.3 果蝇性染色体功能紊乱

雌雄果蝇中常染色质 MSUD 尚未发现。而且，常染色质删除和复制的扩展研究认为这类基因存在的可能性更小。另一方面，着丝粒周边 X 异染色质剔除与 Y 常染色体易位相互作用引起的雄性完全不育得到了很好的证实（见综述McKee 1998）。这种不育不能通过重新加入一个完整的 Y 染色体或一个 X 异染色质所逆转。McKee(1998) 提出的 Y/常染色体易位不育是由一个能够检测未配对或配对不良的性染色体的减数分裂检验点的出现引起的。在以上所描述的各种真核生物 MSUD 的基础上，我们提出了另一种解释——X 或 Y 染色体上未配对序列触发了正常生育所必需的元件沉默。

## 7. 展望与结论

真核细胞染色体曾经被视为大货车，载满基因并被含有端粒“缓冲器”的动粒“发动机”驱动。在酿酒酵母中的早期工作支持一种观点，这些运输车辆关键元件的特性由它们的 DNA 序列决定。然而，结构的分析，如较高等生物的着丝粒，指引我们发现遗传元件由染色质状态编码而不是一个简单的 DNA 序列。“可选择性异染色质”（facultative heterochromatin）被提出用于解释 X 失活，或“惰性”（inertness）蔓延来解释位置效应花斑，这些概念已成为基因组复制、修复、有丝分裂及减数分裂和有关的染色体区域活性机制的所需成分。

在某些方面，表观遗传决定一些关键的功能的现象是很令人吃惊的；DNA 序列乍看似乎决定着染色体遗传的稳定机制。然而，自然界充满了与染色体复制和分离相关的功能“状态变化”的例子，如在哺乳动物和果蝇中新着丝粒，各种系统中的“减数分裂驱动”，种种染色体直接丢失的现象。许多遗传元件与高度稳定的 DNA 重复序列有关，目前原因不明。进化则伴随着大或小规模的基因组交换，例如突变和染色体重排。也许这些令人吃惊的遗传元件的可塑性程度对与 DNA 序列中累积的变化是必须的，可能对进化的持续进行也是必需的。

在一定程度上我们对未来能有所窥见，我们猜测现在所了解的只是冰山一角，染色体功能更多存在于染色质而非 DNA 序列。我们期待着更大的进展。

刘 岩 译
王大亮 校
孙方霖 审校

# 参考文献


Ahmad K. and Golic K.G. 1998. The transmission of fragmented chromosomes in *Drosophila melanogaster*. *Genetics* **148:** 775–792.

Ahmad K. and Henikoff S. 2001. Centromeres are specialized replication domains in heterochromatin. *J. Cell Biol.* **153:** 101–110.

———. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9:** 1191–1200.

Aladjem M.I. and Fanning E. 2004. The replicon revisited: An old model learns new tricks in metazoan chromosomes. *EMBO Rep.* **5:** 686–691.

Amor D.J. and Choo K.H. 2002. Neocentromeres: Role in human disease, evolution, and centromere study. *Am. J. Hum. Genet.* **71:** 695–714.

Baarends W.M., Wassenaar E., van der Laan R., Hoogerbrugge J., Sleddens-Linkels E., Hoeijmakers J.H., de Boer P., and Grootegoed J.A. 2005. Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. *Mol. Cell. Biol.* **25:** 1041–1053.

Bassing C.H., Chua K.F., Sekiguchi J., Suh H., Whitlow S.R., Fleming J.C., Monroe B.C., Ciccone D.N., Yan C., et al. 2002. Increased ionizing radiation sensitivity and genomic instability in the absence of histone H2AX. *Proc. Natl. Acad. Sci.* **99:** 8173–8178.

Baum M., Ngan V., and Clarke L. 1994. The centromeric K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. *Mol. Biol. Cell* **5:** 747–761.

Bean C. J., Schaner C. E., and Kelly W. G. 2004. Meiotic pairing and imprinted X chromatin assembly in *Caenorhabditis elegans*. *Nat. Genet.* **36:** 100–105.

Bennett M. D., Dover G. A., and Riley R. 1974. Meiotic duration in wheat genotypes with or without homoeologous meiotic chromosome pairing. *Proc. R. Soc. Lond. B Biol. Sci.* **187:** 191–207.

Beukeboom L.W. and Werren J.H. 1993. Deletion analysis of the selfish B chromosome, *Paternal Sex Ratio* (PSR), in the parasitic wasp *Nasonia vitripennis*. *Genetics* **133:** 637–648.

Biessmann H. and Mason J.M. 2003. Telomerase-independent mechanisms of telomere elongation. *Cell. Mol. Life Sci.* **60:** 2325–2333.

Black B.E., Foltz D.R., Chakravarthy S., Luger K., Woods V.L., Jr., and Cleveland D.W. 2004. Structural determinants for generating centromeric chromatin. *Nature* **430:** 578–582.

Blasco M.A. 2005. Telomeres and human disease: Ageing, cancer and beyond. *Nat. Rev. Genet.* **6:** 611–622.

Bloom K., Hill A., Kenna M., and Saunders M. 1989. The structure of a primitive kinetochore. *Trends Biochem. Sci.* **14:** 223–227.

Blower M.D. and Karpen G.H. 2001. The role of *Drosophila* CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat. Cell Biol.* **3:** 730–739.

Blower M.D., Sullivan B.A., and Karpen G.H. 2002. Conserved organization of centromeric chromatin in flies and humans. *Dev. Cell* **2:** 319–330.

Carroll C.W. and Straight A.F. 2006. Centromere formation: From epigenetics to self-assembly. *Trends Cell Biol.* **16:** 70–78.

Celeste A., Petersen S., Romanienko P.J., Fernandez-Capetillo O., Chen H.T., Sedelnikova O.A., Reina-San-Martin B., Coppola V., Meffre E., Difilippantonio M.J., et al. 2002. Genomic instability in mice lacking histone H2AX. *Science* **296:** 922–927.

Cenci G., Siriaco G., Raffa G.D., Kellum R., and Gatti M. 2003. The *Drosophila* HOAP protein is required for telomere capping. *Nat. Cell Biol.* **5:** 82–84.

Clarke L., Amstutz H., Fishel B., and Carbon J. 1986. Analysis of centromeric DNA in the fission yeast *Schizosaccharomyces pombe*. *Proc. Natl. Acad. Sci.* **83:** 8253–8257.

Cleveland D.W., Mao Y., and Sullivan K.F. 2003. Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. *Cell* **112:** 407–421.

Collins K.A., Furuyama S., and Biggins S. 2004. Proteolysis contributes to the exclusive centromere localization of the yeast Cse4/CENP-A histone H3 variant. *Curr. Biol.* **14:** 1968–1972.

Cryderman D.E., Morris E., Biessmann H., Elgin S.C., and Wallrath L.L. 1999. Silencing at *Drosophila* telomeres: Nuclear organization and chromatin structure play critical roles. *EMBO J.* **18:** 3724–3735.

Cullen C.F., Brittle A.L., Ito T., and Ohkura H. 2005. The conserved kinase NHK-1 is essential for mitotic progression and unifying acentrosomal meiotic spindles in *Drosophila melanogaster*. *J. Cell Biol.* **171:** 593–602.

Dawe R.K. and Hiatt E.N. 2004. Plant neocentromeres: Fast, focused, and driven. *Chromosome Res.* **12:** 655–669.

De La Fuente R. 2006. Chromatin modifications in the germinal vesicle (GV) of mammalian oocytes. *Dev. Biol.* **292:** 1–12.

de Saint Phalle B. and Sullivan W. 1996. Incomplete sister chromatid separation is the mechanism of programmed chromosome elimination during early *Sciara coprophila* embryogenesis. *Development* **122:** 3775–3784.

Dernburg A.F. 2001. Here, there, and everywhere: Kinetochore function on holocentric chromosomes. *J. Cell Biol.* **153:** F33–38.

Dernburg A.F., Sedat J.W., and Hawley R.S. 1996. Direct evidence of a role for heterochromatin in meiotic chromosome segregation. *Cell* **86:** 135–146.

Dobson S.L. and Tanouye M.A. 1998. Evidence for a genomic imprinting sex determination mechanism in *Nasonia vitripennis* (Hymenoptera; Chalcidoidea). *Genetics* **149:** 233–242.

Donaldson K.M., Lui A., and Karpen G.H. 2002. Modifiers of terminal deficiency-associated position effect variegation in *Drosophila*. *Genetics* **160:** 995–1009.

Dvorak J. and Lukaszewski A.J. 2000. Centromere association is an unlikely mechanism by which the wheat *Ph1* locus regulates metaphase I chromosome pairing between homoeologous chromosomes. *Chromosoma* **109:** 410–414.

Fanti L., Giovinazzo G., Berloco M., and Pimpinelli S. 1998. The heterochromatin protein 1 prevents telomere fusions in *Drosophila*. *Mol. Cell* **2:** 527–538.

Ferguson B.M. and Fangman W.L. 1992. A position effect on the time of replication origin activation in yeast. *Cell* **68:** 333–339.

Fernandez-Capetillo O., Lee A., Nussenzweig M., and Nussenzweig A. 2004. H2AX: The histone guardian of the genome. *DNA Repair* **3:** 959–967.

Ferreri G.C., Liscinsky D.M., Mack J.A., Eldridge M.D., and O'Neill R.J. 2005. Retention of latent centromeres in the mammalian genome. *J. Hered.* **96:** 217–224.

Friedberg E., Walker G., and Siede W. 1995. *DNA repair and mutagenesis*. ASM Press, Washington D.C.

Ganetzky B. 1977. On the components of segregation distortion in *Drosophila melanogaster*. *Genetics* **86:** 321–355.

Garcia-Cao M., O'Sullivan R., Peters A.H., Jenuwein T., and Blasco M.A. 2004. Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat. Genet.* **36:** 94–99.

Goday C. and Ruiz M.F. 2002. Differential acetylation of histones H3 and H4 in paternal and maternal germline chromosomes during development of sciarid flies. *J. Cell Sci.* **115:** 4765–4775.

Grell R.F. 1976. Distributive pairing. In *The genetics and biology of* Drosophila (ed. E. Novitski and M. Ashburner), pp. 436–483. Academic Press, New York.

Griffiths S., Sharp R., Foote T.N., Bertin I., Wanous M., Reader S., Colas I., and Moore G. 2006. Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* **439:** 749–752.

Hall A.E., Kettler G.C., and Preuss D. 2006. Dynamic evolution at pericentromeres. *Genome Res.* **16:** 355–364.

Hassa P.O. and Hottiger M.O. 2005. An epigenetic code for DNA damage repair pathways? *Biochem. Cell Biol.* **83:** 270–285.

Hawley R.S., Irick H., Zitron A.E., Haddox D.A., Lohe A., New C., Whitley M.D., Arbel T., Jang J., McKim K., and Childs G. 1993. There are two mechanisms of achiasmate segregation in *Drosophila* females,

one of which requires heterochromatic homology. *Dev. Genet.* **13:** 440–467.

Heun P., Erhardt S., Blower M.D., Weiss S., Skora A.D., and Karpen G.H. 2006. Mislocalization of the *Drosophila* centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev. Cell* **10:** 303–315.

Huyen Y., Zgheib O., Ditullio R.A., Jr., Gorgoulis V.G., Zacharatos P., Petty T.J., Sheston E.A., Mellert H.S., Stavridi E.S., and Halazonetis T.D. 2004. Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature* **432:** 406–411.

Hynes M.J. and Todd R.B. 2003. Detection of unpaired DNA at meiosis results in RNA-mediated silencing. *Bioessays* **25:** 99–103.

Ivanovska I., Khandan T., Ito T., and Orr-Weaver T.L. 2005. A histone code in meiosis: The histone kinase, NHK-1, is required for proper chromosomal architecture in *Drosophila* oocytes. *Genes Dev.* **19:** 2571–2582.

Joseph I., Jia D., and Lustig A.J. 2005. Ndj1p-dependent epigenetic resetting of telomere size in yeast meiosis. *Curr. Biol.* **15:** 231–237.

Karpen G.H. and Spradling A.C. 1992. Analysis of subtelomeric heterochromatin in the *Drosophila* minichromosome Dp1187 by single P element insertional mutagenesis. *Genetics* **132:** 737–753.

Karpen G.H., Le M.H., and Le H. 1996. Centric heterochromatin and the efficiency of achiasmate disjunction in *Drosophila* female meiosis. *Science* **273:** 118–122.

Kemp B., Boumil R.M., Stewart M.N., and Dawson D.S. 2004. A role for centromere pairing in meiotic chromosome segregation. *Genes Dev.* **18:** 1946–1951.

Kusano A., Staber C., Chan H.Y., and Ganetzky B. 2003. Closing the (Ran)GAP on segregation distortion in *Drosophila*. *Bioessays* **25:** 108–115.

Kutil B.L., Seong K.Y., and Aramayo R. 2003. Unpaired genes do not silence their paired neighbors. *Curr. Genet.* **43:** 425–432.

Lai S.R., Phipps S.M., Liu L., Andrews L.G., and Tollefsbol T.O. 2005. Epigenetic control of telomerase and modes of telomere maintenance in aging and abnormal systems. *Front. Biosci.* **10:** 1779–1796.

Lam A.L., Boivin C.D., Bonney C.F., Rudd M.K., and Sullivan B.A. 2006. Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. *Proc. Natl. Acad. Sci.* **103:** 4186–4191.

Lee D.W., Pratt R.J., McLaughlin M., and Aramayo R. 2003. An argonaute-like protein is required for meiotic silencing. *Genetics* **164:** 821–828.

Lee D.W., Seong K.Y., Pratt R.J., Baker K., and Aramayo R. 2004. Properties of unpaired DNA required for efficient silencing in *Neurospora crassa*. *Genetics* **167:** 131–150.

Lo A.W., Magliano D.J., Sibson M.C., Kalitsis P., Craig J.M., and Choo K.H. 2001. A novel chromatin immunoprecipitation and array (CIA) analysis identifies a 460-kb CENP-A-binding neocentromere DNA. *Genome Res.* **11:** 448–457.

Louis E.J. and Vershinin A.V. 2005. Chromosome ends: Different sequences may provide conserved functions. *Bioessays* **27:** 685–697.

Lue N.F. 2004. Adding to the ends: What makes telomerase processive and how important is it? *Bioessays* **26:** 955–962.

Luo M.C., Dubcovsky J., and Dvorak J. 1996. Recognition of homeology by the wheat *Ph1* locus. *Genetics* **144:** 1195–1203.

Maggert K.A. and Karpen G.H. 2001. Neocentromere formation occurs by an activation mechanism that requires proximity to a functional centromere. *Genetics* **158:** 1615–1628.

Maine E.M., Hauth J., Ratliff T., Vought V.E., She X., and Kelly W.G. 2005. EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during *C. elegans* meiosis. *Curr. Biol.* **15:** 1972–1978.

Malik H.S. and Henikoff S. 2002. Conflict begets complexity: The evolution of centromeres. *Curr. Opin. Genet. Dev.* **12:** 711–718.

McKee B.D. 1998. Pairing sites and the role of chromosome pairing in meiosis and spermatogenesis in male *Drosophila*. *Curr. Top. Dev. Biol.* **37:** 77–115.

Mellone B.G. and Allshire R.C. 2003. Stretching it: Putting the CEN(P-A) in centromere. *Curr. Opin. Genet. Dev.* **13:** 191–198.

Miller K.M. and Cooper J.P. 2003. The telomere protein Taz1 is required to prevent and repair genomic DNA breaks. *Mol. Cell* **11:** 303–313.

Nakaseko Y., Kinoshita N., and Yanagida M. 1987. A novel sequence common to the centromere regions of *Schizosaccharomyces pombe* chromosomes. *Nucleic Acids Res.* **15:** 4705–4715.

Nakatani Y., Ray-Gallet D., Quivy J.P., Tagami H., and Almouzni G. 2004. Two distinct nucleosome assembly pathways: Dependent or independent of DNA synthesis promoted by histone H3.1 and H3.3 complexes. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 273–280.

Nimmo E.R., Pidoux A.L., Perry P.E., and Allshire R.C. 1998. Defective meiosis in telomere-silencing mutants of *Schizosaccharomyces pombe*. *Nature* **392:** 825–828.

Nonaka N., Kitajima T., Yokobayashi S., Xiao G., Yamamoto M., Grewal S.I., and Watanabe Y. 2002. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat. Cell Biol.* **4:** 89–93.

Ogino K., Hirota K., Matsumoto S., Takeda T., Ohta K., Arai K.I., and Masai H. 2006. Hsk1 kinase is required for induction of meiotic dsDNA breaks without involving checkpoint kinases in fission yeast. *Proc. Natl. Acad. Sci.* **103:** 8131–8136.

Pasero P., Bensimon A., and Schwob E. 2002. Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus. *Genes Dev.* **16:** 2479–2484.

Perrini B., Piacentini L., Fanti L., Altieri F., Chichiarelli S., Berloco M., Turano C., Ferraro A., and Pimpinelli S. 2004. HP1 controls telomere capping, telomere elongation, and telomere silencing by two different mechanisms in *Drosophila*. *Mol. Cell* **15:** 467–476.

Pidoux A.L. and Allshire R.C. 2004. Kinetochore and heterochromatin domains of the fission yeast centromere. *Chromosome Res.* **12:** 521–534.

Prasanth S. G., Mendez J., Prasanth K. V., and Stillman B. 2004. Dynamics of pre-replication complex proteins during the cell division cycle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **359:** 7–16.

Pratt R.J., Lee D.W., and Aramaya R. 2004. DNA methylation affects meiotic trans-sensing, not meiotic silencing, in *Neurospora*. *Genetics* **168:** 1925–1935.

Reddy K.C. and Villeneuve A.M. 2004. *C. elegans* HIM-17 links chromatin modification and competence for initiation of meiotic recombination. *Cell* **118:** 439–452.

Rhoades M.M. and Dempsey E. 1966. The effect of abnormal chromosome 10 on preferential segregation and crossing over in maize. *Genetics* **53:** 989–1020.

Rudd M.K., Mays R.W., Schwartz S., and Willard H.F. 2003. Human artificial chromosomes with alpha satellite-based de novo centromeres show increased frequency of nondisjunction and anaphase lag. *Mol. Cell Biol.* **23:** 7689–7697.

Sadaie M., Naito T., and Ishikawa F. 2003. Stable inheritance of telomere chromatin structure and function in the absence of telomeric repeats. *Genes Dev.* **17:** 2271–2282.

Sandler L. and Hiraizumi Y. 1959. Meiotic drive in natural populations of *Drosophila melanogaster*. II. Genetic variation at the *Segregation-distorter* locus. *Proc. Natl. Acad. Sci.* **45:** 1412–1422.

Sharp J.A., Franco A.A., Osley M.A., and Kaufman P.D. 2002. Chromatin assembly factor I and Hir proteins contribute to building functional kinetochores in *S. cerevisiae*. *Genes Dev.* **16:** 85–100.

Shelby R.D., Monier K., and Sullivan K.F. 2000. Chromatin assembly at kinetochores is uncoupled from DNA replication. *J. Cell Biol.* **151:** 1113–1118.

Shiu P.K. and Metzenberg R.L. 2002. Meiotic silencing by unpaired DNA: Properties, regulation and suppression. *Genetics* **161:** 1483–1495.

Steiner N.L. and Clarke L. 1994. A novel epigenetic effect can alter centromere function in fission yeast. *Cell* **79:** 865–874.

Sullivan B. and Karpen G. 2001. Centromere identity in *Drosophila* is not determined in vivo by replication timing. *J. Cell Biol.* **154:** 683–690.

Sullivan B.A. and Karpen G.H. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* **11:** 1076–1083.

Sullivan B.A. and Willard H.F. 1998. Stable dicentric X chromosomes

with two functional centromeres. *Nat. Genet.* **20:** 227–228.

Tsubouchi T. and Roeder G.S. 2005. A synaptonemal complex protein promotes homology-independent centromere coupling. *Science* **308:** 870–873.

Turner J.M., Mahadevaiah S.K., Fernandez-Capetillo O., Nussenzweig A., Xu X., Deng C.X., and Burgoyne P.S. 2005. Silencing of unsynapsed meiotic chromosomes in the mouse. *Nat. Genet.* **37:** 41–47.

Watanabe Y. 2005. Sister chromatid cohesion along arms and at centromeres. *Trends Genet.* **21:** 405–412.

Weinreich M., Palacios DeBeer M.A., and Fox C.A. 2004. The activities of eukaryotic replication origins in chromatin. *Biochim. Biophys. Acta* **1677:** 142–157.

Werren J.H., Nur U., and Eickbush D. 1987. An extrachromosomal factor causing loss of paternal chromosomes. *Nature* **327:** 75–76.

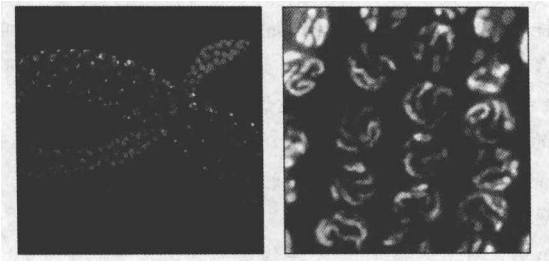
Wu H.Y. and Burgess S.M. 2006. Ndj1, a telomere-associated protein, promotes meiotic recombination in budding yeast. *Mol. Cell Biol.* **26:** 3683–3694.

Yan H., Jin W., Nagaki K., Tian S., Ouyang S., Buell C.R., Talbert P.B., Henikoff S., and Jiang J. 2005. Transcription and histone modifications in the recombination-free region spanning a rice centromere. *Plant Cell* **17:** 3227–3238.

Yoda K., Ando S., Morishita S., Houmura K., Hashimoto K., Takeyasu K., and Okazaki T. 2000. Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. *Proc. Natl. Acad. Sci.* **97:** 7266–7271.

Yu H.G., Hiatt E.N., Chan A., Sweeney M., and Dawe R.K. 1997. Neocentromere-mediated chromosome movement in maize. *J. Cell Biol.* **139:** 831–840.

# 第 15 章



# 秀丽隐杆线虫 X 染色体的表观遗传学调控

Susan Strome[1] and William G. Kelly[2]

[1]*Department of Biology, Indiana University, Bloomington, Indiana 47405*
[2]*Biology Department, Emory University, Atlanta, Georiga 30322*

## 本章目录

1. 秀丽隐杆线虫性染色体的不均衡性
2. DCC 类似于致密因子复合物
3. X∶A 比值的确定
4. DCC 的募集与蔓延
5. DCC 效应：X 连锁基因和常染色体基因 *her-1* 的下调
6. 生殖系中 X 染色体的调节
7. 生殖系发育和 X 染色体沉默
8. 雄性个体中的单一 X 染色体呈现出异染色质特征
9. 生殖系中 MSUD 对基因表达模式的影响
10. MES 组蛋白修饰酶对 X 染色体沉默的调控
11. 精细胞 X 染色体的印记化
12. 总结

参考文献

## 概要

土壤线虫之秀丽隐杆线虫的性别决定机制取决于X染色体与常染色体组的比值（X∶A比值）。有两条X染色体（X∶A比值为1）的二倍体线虫发育成雌雄同体，而那些只有一条X染色体（X∶A比值为0.5）的线虫发育成雄虫。如果不被调节的话，不同性别中X染色体剂量的不同会导致X连锁基因表达水平的不同，进而导致一种性别的死亡。本章讨论在两种性别的体细胞组织和生殖系中对不同数量X染色体进行基因表达调节的各种表观遗传策略，并将线虫使用的策略与其他物种解决类似的X染色体剂量差异所采用的策略进行比较。

在线虫的体细胞组织中，对不同性别中X连锁基因表达水平的平衡，或称“剂量补偿效应”，是通过对XX个体中两条X染色体的基因表达同时下调一半来实现的。这种下调通过剂量补偿复合物（DCC）来完成，该复合物类似于致密因子复合物。后者参与了细胞内所有染色体的致密化过程，这一过程在核分裂产生两个子代细胞时有助于染色体组的分离。因此，DCC可能通过部分致密化X染色体实现对X连锁基因的抑制。线虫剂量补偿效应的关键问题在于X∶A比值如何被确定，DCC如何在XX线虫中被特异性组装，DCC如何被定位到X染色体上以及DCC如何精确地将X连锁基因的表达水平下调一半。这几个方面的机制研究都取得了重大的进展。

线虫所使用的剂量补偿效应机制有别于哺乳动物和果蝇（Drosophila）。哺乳动物通过完全沉默XX个体中的一条X染色体来实现剂量补偿。而果蝇是通过XY个体中单一X染色体上基因表达的上调来实现的。这种机制的多样性可能反映了，具有不同数量X染色体的动物为平衡X连锁基因表达这一特殊任务，从多种可供选择的已有系统中作出了各自的选择。

在线虫的生殖系组织（如生殖细胞）中存在着一种更为极端的X连锁基因表达的调节模式：雄性个体中唯一的X染色体和雌雄同体中的两条X染色体都被沉默了。在两种性别的生殖系细胞中，X染色体上均缺少存在于活跃表达染色质的组蛋白修饰。这至少部分地由MES蛋白调控。此外，在雄性个体中，每个生殖核中单一X染色体上都获得了与异染色质沉默相关的组蛋白修饰。这种沉默依赖于雄性个体中X染色体在减数分裂时的非配对状态。表达于生殖系的基因中，分布于X染色体上的数量惊人的低。一种被偏好的观点是，在雄性个体减数分裂的生殖细胞中，单一X染色体的强烈抑制导致了X染色体上罕有在生殖系表达的基因，也导致了在XX个体生殖系中对X染色体进行沉默的必要性。因此，秀丽隐杆线虫生殖系代表了一种有趣的研究案例，可以用于研究不同性别中染色体的不均衡是如何导致染色体状态的表观遗传学调控并决定组织中的基因表达谱。

## 1. 秀丽隐杆线虫性染色体的不均衡性

秀丽隐杆线虫以两种性别存在，它们在遗传上通过 X 染色体数量的不同而被区分开：XX 线虫是雌雄同体而 XO 线虫是雄虫。在线虫中不存在性别特异的染色体，如 Y 染色体。雌雄同体和雄虫表现出多种性别特异的解剖学特征，并且有不同的生殖系程序（图 15-1）。这些性别间的显著差异始于胚胎早期，是对 X 染色体数量计数并作出反应的结果（Nigon 1951；综述可参见 Meyer 2000）。性染色体数量或倍数的简单差异是如何导致发育程序的这种显著差异呢？一个重要的观点是：每一个 *C. elegans* 细胞不仅需要确定自己的 X 染色体数量，还要确定其常染色体组的数量。事实上是 X：A 的比值决定了线虫的性别。有两条 X 染色体（X：A 比值为 1）的双倍体线虫发育成雌雄同体，而那些只有一条 X 染色体（X：A 比值为 0.5）的发育成雄虫。人们对于正确应答 X：A 比值的机制细节进行了精细的剖析，详细描述见下文。

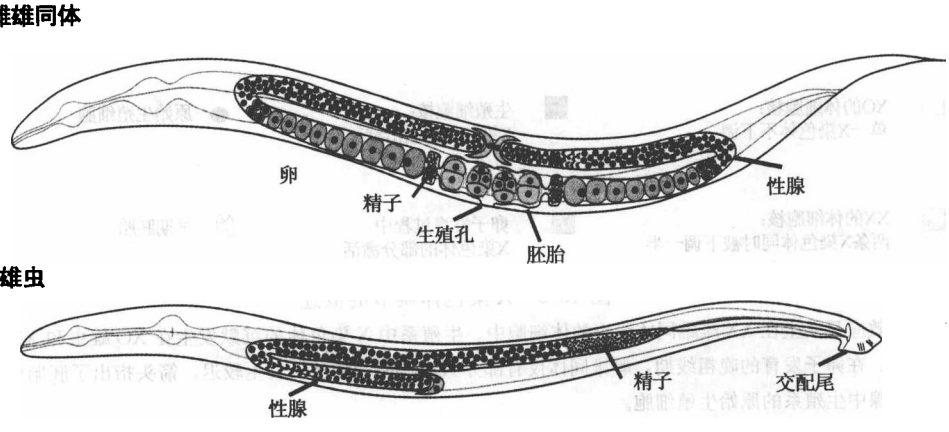



图 15-1　秀丽隐杆线虫 XX 和 XO 解剖图

自然界中的秀丽隐杆线虫以两种性别存在：XX 雌雄同体和 XO 雄虫。雌雄同体线虫和雄虫表现出许多性别特有的解剖学性征，最特别的是雄虫有一条用于交配的尾而雌雄同体线虫的腹部表面存在一个阴户用于接受雄虫精子和产卵。它们的生殖系发育也不相同。雌雄同体线虫中双臂型的性腺最初能产生精子而后在成体时期产生卵母细胞。雄虫中单臂型的性腺能持续产生精子（经允许改编自 Hansen et al. 2004[© Elsevier]）。

性别间 X 染色体剂量的差异如果不被调节的话，可能会导致 X 连锁基因表达水平的不同和一种性别的致死。有趣的是，体细胞和生殖细胞进化出了应对 X 剂量问题的不同机制（图 15-2）。生殖系和体细胞世系在胚胎发育的 16～24 细胞时期就已相互分离。大约在 30 细胞时期，体细胞世系开启了一个被称为“剂量补偿效应”的程序。这建立起了一个在 XX 线虫中的两条 X 染色体同时被下调一半的模式。与之相反，正如第 16 章和第 17 章讨论的那样，哺乳动物通过全面沉默 XX 性别中的一条 X 染色体而实现剂量补偿，而果蝇则通过上调 XY 个体中单一 X 染色体基因的表达来完成剂量补偿。在秀丽隐杆线虫生殖系组织中存在着一种更为极端的 X 连锁基因表达的调节：雄虫中单一 X 和雌雄同体中的两条 X 染色体都被全面沉默。本章的主题就是探讨在体细胞组

织中实现剂量补偿效应和生殖系中实现 X 染色体沉默的表观遗传学机制。

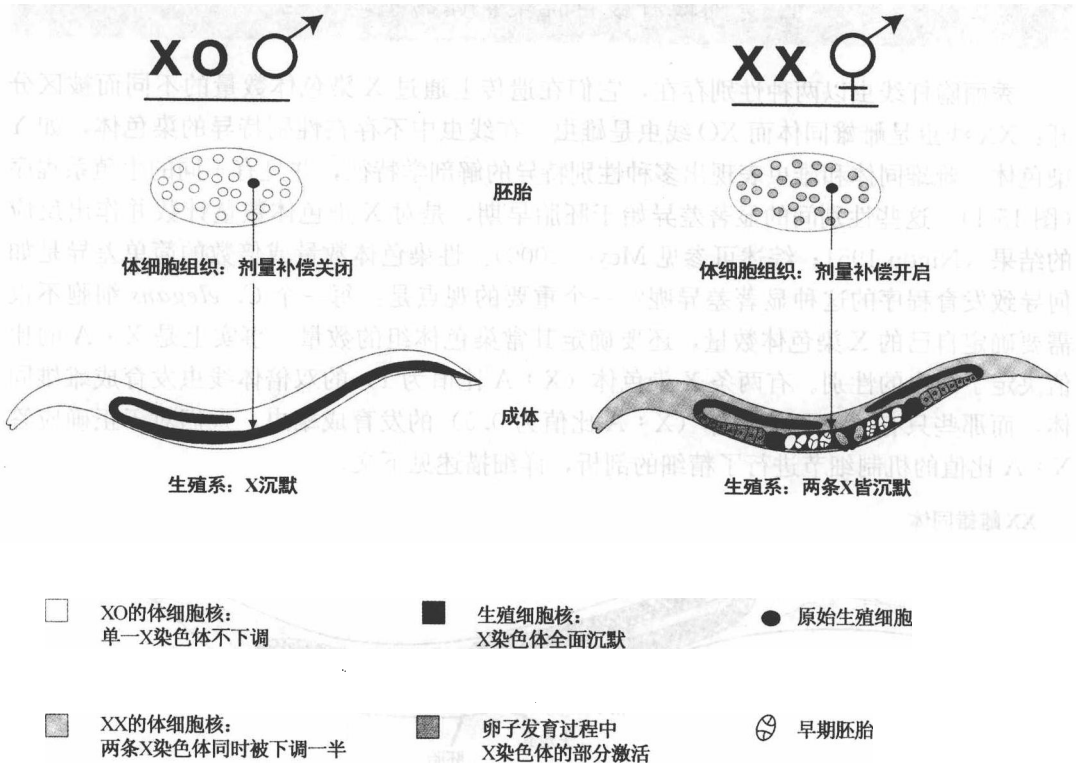



图 15-2 X 染色体调节的概述

剂量补偿效应只发生在 XX 雌雄同体线虫的体细胞中。生殖系中 X 染色体的沉默发生在 XO 雄虫和 XX 雌雄同体线虫中。在卵子发育的晚粗线期，雌雄同体仅有部分 X 连锁基因激活且发生较迟。箭头指出了胚胎中能发育成成体性腺中生殖系的原始生殖细胞。

## 2. DCC 类似于致密因子复合物

要想了解剂量补偿复合物（DCC）的装配和组成，就需要先简单了解一下调节性别决定和剂量补偿通路中的一些首要基因（综述请见 Meyer 1997）。作为这个通路中的首要基因，*xol-1*（*xol* 指 XO 致死）被认为是一个主要的控制基因，因为它的活性是由 X：A的比值决定，并进而决定这一通路是导致雄性还是雌雄同体发育。XOL-1 是三种 *sdc* 基因（*sdc* 指性别决定和剂量补偿缺陷，*sex determination and dosage compensation defective*）的负调节因子。*sdc* 基因编码 DCC 的组分并调节 *her-1*（*her* 指 XO 线虫的雌雄同体化，*hermaphroditization*）性别决定基因。在 XO 胚胎中，X：A 比值为 0.5，导致 XOL-1 蛋白的高水平和 SDC 蛋白的低水平；DCC 不组装，剂量补偿就不能发生，性别决定基因 *her-1* 得以表达导致雄性性别发育。在 XX 胚胎中，X：A 比值为 1，导致 XOL-1 蛋白的低水平和 SDC 的高水平；DCC 得以组装，发生剂量补偿效应，继而抑制 *her-1* 的表达，导致雌雄同体性别发育。DCC 组装特别值得思考的是，作为对 X：A 比

值为 1 的应答，SDC 蛋白在 XX 胚胎中特异性表达。

在 SDC 蛋白（SDC-1、SDC-2 和 SDC-3）以外，DCC 还包括一系列 DPY(DPY 指胖的，dumpy）蛋白（DPY-21、DPY-26、DPY-27、DPY-28 和 DPY-30）和 MIX-1（MIX 指有丝分裂和 X 染色体相关，mitosis and X-associated）蛋白（表 15-1）（综述见 Meyer 2005）。对线虫剂量补偿机制的重要理解来自于秀丽隐杆线虫 DCC 的一部分类似于 13S 致密因子复合物的发现（表 15-1 和图 15-3）（综述见 Meyer 2005）。致密因子复合物在物种间保守，并对在有丝分裂和减数分裂过程中，正确进行染色体致密化和分离至关重要（综述见 Hirano 2002）。MIX-1、DPY-27、DPY-28 和 DPY-26 尤其与致密因子复合物必需组分相同源；MIX-1 也是有丝分裂和减数分裂过程中正常致密因子复合物的基本组分。SDC 蛋白和 DPY-30 与致密因子复合物的已知亚基不相近。目前的观点是 DCC 复合物起源于某种致密因子复合物，特异性地定位于并下调 X 染色体基因，并可能通过导致一定程度的染色质致密化来实现其效应。

**表 15-1　调控 X 染色体基因表达的组分**

| 蛋白质 | 复合物 | 同源物和(或)保守功能域 |
|---|---|---|
| SDC-1 | DCC | $C_2H_2$ 锌指功能域 |
| SDC-2 | DCC | 新蛋白 |
| SDC-3 | DCC | $C_2H_2$ 锌指功能域和类似于肌浆球蛋白的 ATP 结合功能域 |
| DPY-21 | DCC | 保守蛋白;无明显功能域 |
| DPY-26 | DCC | 致密因子亚基 XCAP-H |
| DPY-27 | DCC | 致密因子亚基 SMC-4/XCAP-C |
| DPY-28 | DCC | 致密因子亚基 XCAP-D2/Cnd1/Ycs4p |
| DPY-30 | DCC | 保守蛋白;无明显功能域 |
| MIX-1 | DCC | 致密因子亚基 SMC-2/XCAP-E |
| MES-2 | MES-2/3/6 | PRC2 亚基 E(Z)/EZH2;SET 功能域 |
| MES-3 | MES-2/3/6 | 新蛋白 |
| MES-6 | MES-2/3/6 | PRC2 亚基 ESC/EED;WD40 功能域 |
| MES-4 | 未知 | NSD1;PHD 结构域和 SET 功能域 |

为什么 DCC 只在 XX 胚胎中组装呢？奇怪的是，大部分的 DCC 组分在 XX 和 XO 胚胎中都是通过卵细胞由母源提供的。SDC-2 是 DCC 组装的关键调节因子（Dawes et al. 1999）。SDC-2 不由母源提供而只在 XX 胚胎中生成（图 15-4），在 XX 胚胎中，SDC-2 和 SDC-3，DPY-30 一起将其余的 DCC 亚基募集到 X 染色体上。事实上，诱导 XO 胚胎中 SDC-2 的表达足以导致单一 X 染色体上 DCC 的组装并进而引发能杀死胚胎的剂量补偿效应。因此 SDC-2 指导了具有其他细胞学功能的 DCC 组分在 X 染色体的特异性地募集，共同发挥其剂量补偿和性别决定的活性。

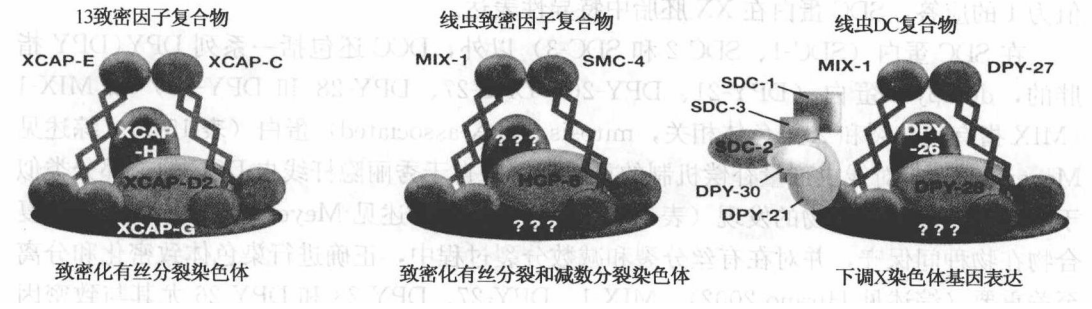



图 15-3 剂量补偿复合物和致密因子复合物

线虫的 DCC 类似于致密因子复合物，后者能在核分裂过程中使染色体致密化。特别的是，DCC 包含许多类似于 XCAP(爪蟾染色体相关多肽，*Xenopus* chromosome-associated polypeptide）的亚基，XCAP 是最初发现于爪蟾的 13S 致密因子复合物的几个亚基。线虫自身的致密因子复合物与同源的爪蟾中功能性的 13S 复合物大体相似。SDC 蛋白和 DPY-30 不同于已知的致密因子复合物亚基；它们在定位 DCC 复合物到 X 染色体的过程中发挥作用（来自于 Meyer 2005）。

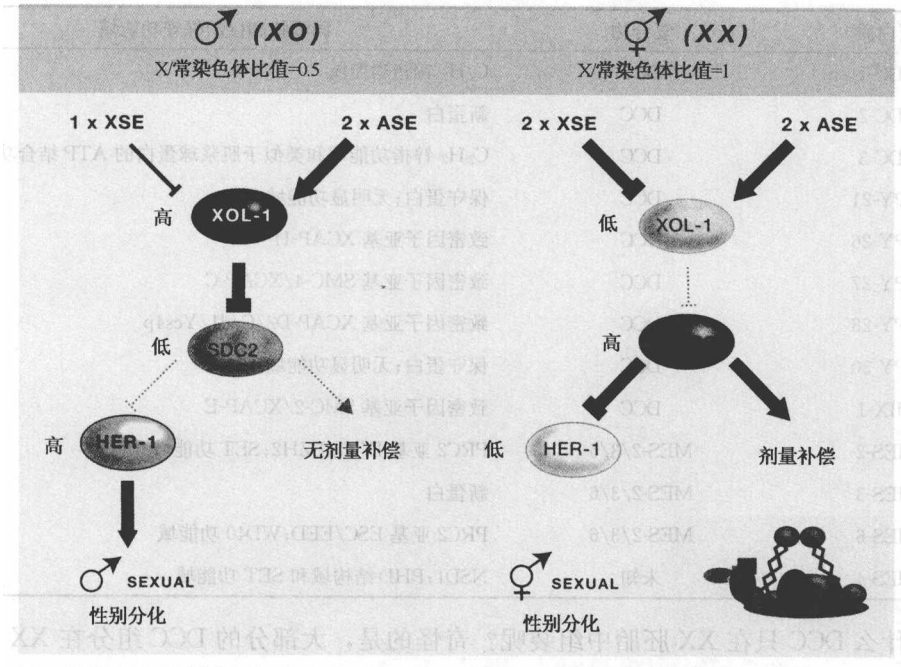



图 15-4 秀丽隐杆线虫中性别决定和剂量补偿效应的通路

本图强调了 X 信号元件（XSE）和常染色体信号元件（ASE）在调节 *XOL-1* 水平时可能发挥的作用，以及随后的性别分化和剂量补偿复合物的组装，后者在 XX 雌雄同体中双倍下调 X 连锁基因的表达水平。

## 3. X：A 比值的确定

线虫细胞是如何计数 X 染色体和常染色体，并在 X：A 的比值为 1 时使用效量补偿效应的呢？X 染色体的四个被称为 X 信号元件（XSE）的小区域被鉴定参与了确认

X：A比值中的分子部分（即被除数，译者注）。在其中两个区域中，相关基因已被鉴定出：*sex*-1 和 *fox*-1(*sex* 指 X 染色体上的信号元件，*signal element on X*；*fox* 指 X 染色体上的雌性化基因，*feminizing gene on X*）（Carmi et al. 1998；Skipper et al. 1999）。二者的基因产物均在胚胎发育的重要时期抑制 *xol-1* 基因的表达，后者为性决定和剂量补偿效应通路的最上游基因（图 15-4）。因为相对于 XO 胚胎，XX 胚胎产生约两倍的 SEX-1 和 FOX-1 蛋白，所以 XX 胚胎中 *xol-1* 的表达水平远远低于 XO 胚胎。SEX-1 是一个抑制 *xol-1* 转录的核激素受体（Carmi et al. 1998）。FOX-1 则是一种 RNA 结合蛋白，它通过某种转录后机制降低 *xol-1* 的 mRNA 水平（Skipper et al. 1999）。因此，X 染色体的剂量与 XSE 表达的抑制 *xol-1* 表达的 X 连锁因子的剂量相当。到目前为止，只有一个常染色体信号元件（ASE）被确定参与了确认 X：A 比值中的分母部分（即除数，译者注）（Powell et al. 2005）。该基因，*sea-1*，编码一个能激活 *xol-1* 转录的 T 盒转录因子。

因此，常染色体剂量（ASE 剂量）通过对重要开关基因 *xol-1* 的拮抗作用来平衡 X 剂量（XSE 剂量）效应（图 15-4）。这种现象在两种性别中是如何发生的呢？最为合理的猜想是：二倍体 XX 胚胎产生了双倍剂量的 *xol-1* 抑制因子 XSE，从而压制 ASE 的激活作用。这样就能够保持 XOL-1 的低水平，并导致雌雄同体的发育和剂量补偿效应的发生。相反的，二倍体 XO 胚胎产生单倍剂量的 XSE 抑制因子，不足以抵消 ASE 的激活作用。高水平的 XOL-1 导致雄虫的发育，不激活剂量补偿效应。XSE 和 ASE 将 X：A比值从 0.5 到 1 的转换翻译成性别转变的开关以及是否发生剂量补偿效应。

## 4. DCC 的募集与蔓延

研究的焦点是发现参与募集 DCC 的 X 染色体特征性元件。值得注意的是在哺乳动物和果蝇中，DCC 同源物是通过与 X 特异的非编码 RNA 和 X 染色体序列元件（参见第 16、17 章）的共同作用被定位到 X 染色体上的。到目前为止，没有证据表明非编码 RNA 在秀丽隐杆线虫的剂量补偿效应中起作用。一个精妙的实验方法被用于研究线虫 DCC 是分别被募集到许多位点上，还是仅有的一个或几个位点再蔓延至邻近的染色体区域（Csankovszki et al. 2004；Meyer 2005）。该实验对具有 X 染色体不同区域倍增副本的线虫种系进行了 DCC 组分的染色分析。DCC 与某个倍增副本的结合意味着该倍增副本及其相应的模板 X 染色体区域含有一个 DCC 募集位点。若 DCC 不与倍增副本结合，而与其相应的模板 X 染色体结合，则意味着该区域中无 DCC 募集位点而需要 DCC 从邻近区域蔓延而来。这些实验至少证明了 13 个能独立募集 DCC 的区域，并为 DCC 沿 X 染色体蔓延提供了证据（图 15-5）（Csankovszki et al. 2004）。一些区域募集能力强而另一些较弱，表明在 X 染色体上或者是募集位点的存在数量不同，或者是不同位点具有募集或促进 DCC 沿 X 染色体蔓延能力的差异。随着目标区域的逐步缩小，已将一个区域的 DCC 募集位点缩小至 33bp(Meyer 2005)。该 33bp 序列在基因组中是单一的。定义其他的最小募集位点是一项优先任务。另一个关键的问题是要解释 DCC 是如何从起始募集位点蔓延的。蔓延可能是由 DCC 间的协同作用或是靠染色质的局部修饰

产生一个更利于DCC结合的自我增强通路来介导的，这已经被裂殖酵母中异染色质的蔓延所证实（见第8章）。

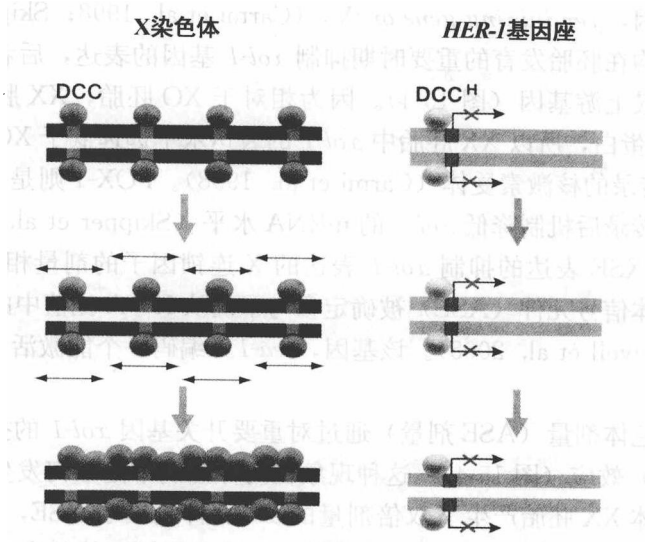



图15-5 DCC介导的X染色体和 *her-1* 基因座的下调

DCC能使X染色体基因的表达减半并20倍下调雌雄同体线虫体细胞组织中 *her-1* 的表达。在XX线虫中DCC发生了组装。它被募集到识别元件上并沿着两条X染色体蔓延，进而两倍下调若干X连锁基因的表达。它也能结合到常染色体 *her-1* 基因的上游区域并20倍下调其表达。$DCC^{H}$ 指与 *her-1* 结合的复合物缺少DPY-21蛋白（经允许改编自Alekseyenko和Kuroda 2004；由Katherine Sutliff说明［© AAAS］）。

## 5. DCC效应：X连锁基因和常染色体基因 *her-1* 的下调

在线虫和果蝇中，调节染色质的结构以实现对X连锁基因表达进行细微的一半水平调节似乎在机制上颇具挑战性。在果蝇中，基因表达的调节伴随着特殊的组蛋白修饰和X染色体上连接组蛋白的消失（见第16章）。在线虫的体细胞中，两条发生剂量补偿效应的X染色体的组蛋白修饰水平或种类相对于雌雄同体的常染色体来说并没有明显的差别。哺乳动物、果蝇和线虫似乎采用了不同的已经存在的系统来实现调节X染色体基因表达的特殊作用。哺乳动物采用了基于异染色质沉默的方式来抑制XX性别中两条X染色体中的一条。果蝇利用染色质修饰机器来改变XY性别中单一X染色体的状态，导致基因的上调表达。秀丽新小杆线虫采用一个用于有丝分裂和减数分裂的染色体浓缩机制来下调XX性别中的两条X染色体。线虫中精确的下调机制以及下调是如何限于一半的仍是一个关键问题。

鉴于DCC类似于13S致密因子复合物，一个可能的基因表达下调的控制机制是通过DCC介导的染色质致密化。这可能限制RNA聚合酶或转录因子接近启动子区域，妨碍RNA聚合酶沿转录单位的行进，亦或减慢每个基因的转录重新起始速度。抑制机制的线索可能能够在单一的常染色体 *her-1* 基因座上找到，因为它是唯一已知的位于常染色体的DCC靶点并且表现出更为极端的受DCC介导的约20倍的表达下调效应（图

15-6)（Dawes et al. 1999；Chu et al. 2002）。*Her-1* 的抑制促进雌雄同体性别的发育（图 15-4）。因此，DCC 能够对 *her-1* 基因实现 10 倍于对 X 染色体的抑制效应。X 染色体结合的 DCC 和 *her-1* 结合的复合物的一个已知区别是 DPY-21。DPY-21 存在于 X 染色体上而不存在于 *her-1* 基因座上（Yonker and Meyer 2003）。

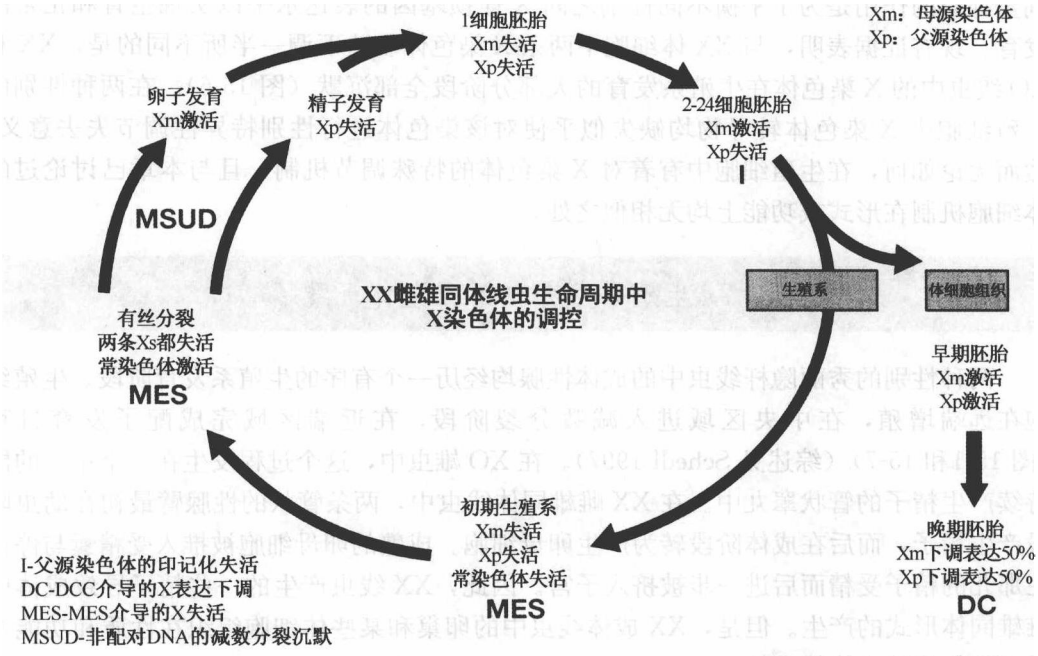



图 15-6　XX 雌雄同体线虫生命周期中 X 染色体的调控

X 染色体在不同阶段和不同组织中被不同的机制所调控：早期胚胎中，父源 X 染色体的印记化 X 失活（I）；30 细胞期的体细胞组织、晚期胚胎和成虫中的剂量补偿（DC）和 MES 介导的生殖系沉默。非配对 DNA 的减数分裂沉默同时也发生在生殖系中；MSUD 可沉默 XO 雄虫中唯一的 X 染色体。

13S 致密因子复合物对于在全基因组范围调节染色体结构进而促进染色体的分离是必须的。对致密因子复合物的使用和改变形成了 DCC，使得它的作用更为精细并且仅对单一染色体起调节作用。具有 DPY-21 的 DCC 轻微地抑制了 X 染色体上许多基因的转录，而没有 DPY-21 的 DCC 却能强烈地抑制 her-1 基因座的转录。今后关于不同致密因子/DCC 复合物是如何发挥不同功能的机制性研究令人期待。

## 6. 生殖系中 X 染色体的调节

在秀丽隐杆线虫中，由于体细胞中 X 染色体的调节模式（例如，通过 DCC）在生殖细胞中不起作用，因此生殖系利用不同于体细胞组织的策略来调节 X 染色体基因的表达。一些体细胞 DCC 的基本组成成分不在生殖系中表达（如 SDC-2）；其他的在生殖细胞中表达，但分布于所有的染色体上（如 MIX-1，DPY-26 和 DPY-28）。如上所说，MIX-1 是秀丽隐杆线虫中 DCC 和常规的致密因子复合物的共同组分。现有的观点认为

致密因子类似复合物扮演了一个特殊的 X 染色体特有的角色，但是它似乎只在体细胞组织中发生作用。在生殖系中，这些被表达的 DCC 蛋白可能在有丝分裂和减数分裂中染色体的组织和分离方面发挥更为普遍性的作用。

生殖系中 DCC 组分的缺失使人们开始怀疑这种组织中是否存在一种剂量补偿模式。剂量补偿的作用是为了平衡不同性别之间 X 连锁基因的表达水平以实现生育和正常的发育。现有证据表明，与 XX 体细胞中两条 X 染色体表达下调一半所不同的是，XX 和 XO 线虫中的 X 染色体在生殖系发育的大部分阶段全部沉默（图 15-6）。在两种性别的生殖细胞中 X 染色体转寻物均缺失似乎使对该染色体进行性别特异性调节失去意义。然而无论如何，在生殖细胞中有着对 X 染色体的特殊调节机制，且与本章已讨论过的体细胞机制在形式或功能上均无相似之处。

## 7. 生殖系发育和 X 染色体沉默

两种性别的秀丽隐杆线虫中的成体性腺均经历一个有序的生殖系发育阶段。生殖细胞在远端增殖，在中央区域进入减数分裂阶段，在近端区域完成配子发育过程（图 15-1和 15-7）（综述见 Schedl 1997）。在 XO 雄虫中，这个过程发生在一个单一的能持续产生精子的管状睾丸中。在 XX 雌雄同体线虫中，两条管状的性腺臂最初在幼虫阶段产生精子，而后在成体阶段转为产生卵母细胞。成熟的卵母细胞被推入受精囊与停留在那儿的精子受精而后进一步被挤入子宫。因此，XX 线虫产生的一些精子导致成体中雌雄同体形式的产生。但是，XX 成体线虫中的卵巢和某些体细胞组织在性质和功能本质上都是雌性个体特有的。

在 XX 雌雄同体线虫和 XO 雄虫中，生殖系的 X 染色体都是沉默的。这种 X 染色体沉默的特征是存在抑制型的组蛋白标记，如 H3K27 甲基化；雌雄同体线虫生殖细胞中的 X 染色体尤其富含 H3K27me3（Bender et al. 2004）。通过免疫荧光分析发现沉默的 X 染色体上也缺少通常与活跃染色质相关的组蛋白乙酰化以及与活跃转录区域相关的 H3K4 甲基化（图 15-7）（Kelly et al. 2002）。秀丽隐杆线虫生殖系中 X 染色体的沉默主要是以下两个机制来实现的（图 15-6）。第一种是通过非配对 DNA 的减数分裂沉默（MSUD），这包括了 XO 线虫中单一 X 染色体的沉默。第二种是通过由 MES 蛋白介导的组蛋白修饰；这种沉默机制在 XX 线虫中被广泛研究，但事实上它也在 XO 雄虫中起作用。在早期胚胎中，父源 X 染色体的沉默在受精之后持续若干细胞周期，被称为印记化的 X 沉默（图 15-6）。

## 8. 雄性个体中的单一 X 染色体呈现出异染色质特征

生殖系中 X 染色体不同于常染色体的最初证据来自于细胞学观察。雄虫中单一 X 染色体在减数分裂的粗线期呈现出一种特征性的结构：它致密化形成一个类似于在哺乳动物减数分裂时可见的 XY“性体”的球状结构（Goldstein and Slaton 1982；Handel 2004）。常染色体的致密化则较晚，发生在减数分裂初期，将要起始精子发生时。在

XX 雌雄同体线虫和性转换的 XX 雄虫的精子减数分裂阶段同样可见提前发生的 X 染色体致密化，表明 X 染色体提前致密化是应答于生殖细胞的性征而非 X 染色体的倍数或配对状态。因此，在 XX 雌雄同体线虫中，将发育成卵的生殖细胞不会发生 X 染色体的提前致密化。

除了 X 染色体的提前致密化，XO 雄虫中单一 X 染色体上有瞬时性的 H3K9me2 的明显富集，这种富集出现在粗线期而在双线期消失（图 15-7）（Kelly et al. 2002）。H3K9me2 的富集并不在 XX 精子发生时、雌雄同体线虫或性转换的 XX 雄虫中发生。但是，在性转换的 XO 雌雄同体线虫中，X 染色体上 H3K9me2 的富集发生在卵母细胞形成的粗线期，与 XO 雄虫中发生的 H3K9me2 的积累相类似（Bean et al. 2004）。因此，在 XO 减数分裂过程中，X 染色体上特殊的异染色质特征的获得似乎取决于它的未配对状态而非其将成为的生殖细胞的性征。H3K9me2 定位于未配对 DNA 不仅局限于 X 染色体序列，它也发生在未配对的常染色体片段和易位片段上（Bean et al. 2004）。

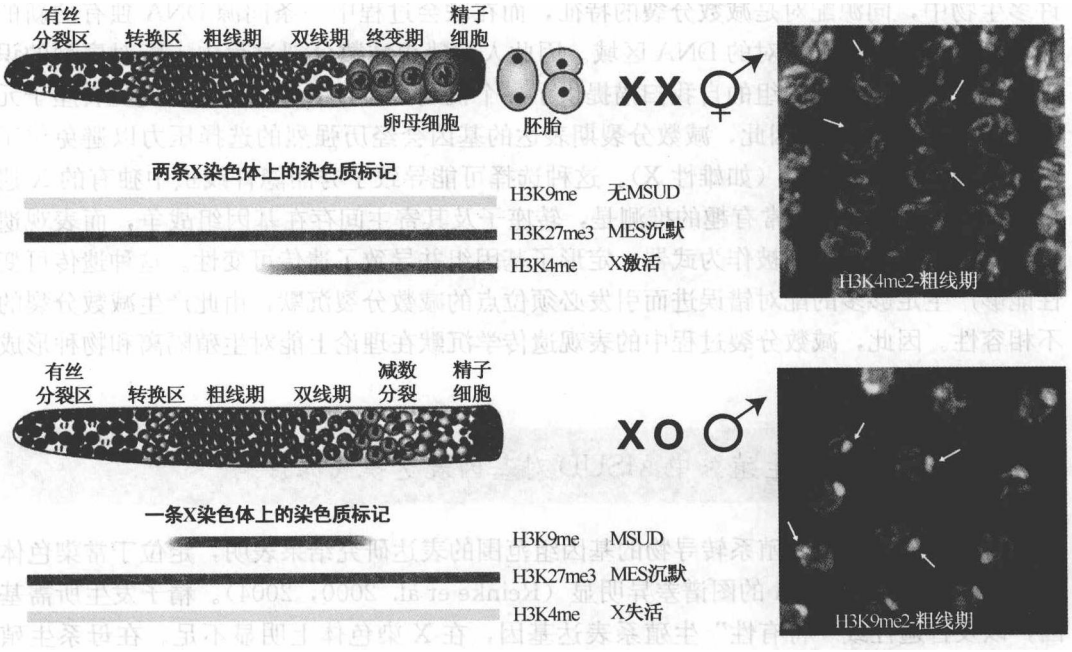



图 15-7　生殖细胞发育过程中 X 染色体的表观遗传调控

在两种性别中，生殖细胞进行有丝分裂（左）并在转换区进入减数分裂，然后发育至减数分裂 I 早期；即将形成精子的细胞都在性腺中完成减数分裂。在雌雄同体线虫中，将形成卵母细胞的细胞在性腺中发育至减数分裂前期然后在排卵和受精之后完成减数分裂。存在于生殖细胞 X 染色体上的各种组蛋白修饰以红色（表示抑制型修饰）和绿色（表示激活型修饰）横条表示。如右图所示，特定的组蛋白修饰抗体显示，在生殖核中 X 染色体的标记与常染色体不同且被沉默。当活跃表达染色质标记时，H3K4me2（绿色）被排除于 XX 粗线期核 X 染色体之外。当异染色质标记时，H3K9me2（绿色）在 XO 粗线期核的 X 染色体上富集。DNA 被染为红色。箭头指出每一个图像中的有代表性的 X 染色体。

减数分裂中未配对 DNA 的靶向抑制并非秀丽隐杆线虫所特有。相似的，未配对 DNA 的识别和抑制也发生在其他生物的减数分裂中，包括链孢霉和小鼠，这被称为未

配对 DNA 的减数分裂沉默（MSUD）（Shiu et al. 2001；Baarends et al. 2005；Turner et al. 2005）。如雄性小鼠减数分裂时联会很差的 XY“性小体”也因其非配对状态而富集 H3K9me2(Cowell et al. 2002；Turner et al. 2005)。链孢霉中发生的减数分裂沉默需要在 RNA 干扰（RNAi）过程中具有保守功能的一些蛋白活性，其中包括一个 RNA 介导的 RNA 聚合酶（RdRP），一个 Argonaute 相关蛋白（RNA 诱导的沉默复合物，RISC 的保守成分）和 Dicer 核酸酶（综述见 Nakayashiki 2005；第 8 章）。秀丽隐杆线虫中未配对 DNA 上 H3K9me2 的富集需要一个生殖系特有的 RdRP：EGO-1，但不需要 Dicer(Maine et al. 2005)。这表明虽然减数分裂沉默的现象（未配对 DNA 的抑制）是保守的，但不同生物中实现沉默的机制却不尽相同。

与 XO 雄虫减数分裂时所发生的现象不同的是，在 XX 精子发生或 XX 卵子发生过程中 X 染色体并不发生 H3K9me2 的富集。这个差异可能是由于雌雄同体线虫中发生了 X 染色体的完全联会。那么为什么未配对 DNA 的沉默会是有性生殖的保守特征呢？在许多生物中，同源配对是减数分裂的特征，而在联会过程中一条同源 DNA 独有的新的插入可能会被视作未配对的 DNA 区域。因此人们猜想减数分裂过程中未配对序列的识别和沉默为双倍体基因组的自我扫描提供了一个机制，并且能够保护机体免受转座子元件的入侵（或扩散）。因此，减数分裂期表达的基因会经历强烈的选择压力以避免位于一个未配对的染色体上（如雄性 X）。这种选择可能导致了秀丽隐杆线虫中独有的 X 遗传图谱。因此，一种非常有趣的推测是，转座子及其寄主间存在基因组战争，而表观遗传学机制在这场战争中被作为武器，定形了基因组并导致了遗传可变性。这种遗传可变性能够产生足够多的配对错误进而引发必须位点的减数分裂沉默，由此产生减数分裂的不相容性。因此，减数分裂过程中的表观遗传学沉默在理论上能对生殖隔离和物种形成作出贡献。

## 9. 生殖系中 MSUD 对基因表达模式的影响

对秀丽隐杆线虫生殖系转寻物的基因组范围的表达研究结果表明，定位于常染色体和 X 染色体上的 mRNA 的图谱差异明显（Reinke et al. 2000，2004）。精子发生所需基因，以及普遍性或“特有性”生殖系表达基因，在 X 染色体上明显不足。在母系生殖系大量表达的基因在 X 染色体上同样不足，但其程度不及精子发生或“特有性”生殖系基因。换句话说，生殖细胞发育和功能所需要的基因在极大程度上不存在于 X 染色体上。的确，X 染色体上大部分与卵子发生相关的基因都表达于减数分裂晚期，这与该时期发生的 X 染色体激活相吻合（图 15-6、图 15-7）。在粗线期末期“激活性”组蛋白修饰的富集为减数分裂晚期的激活机制提供了证据，这些组蛋白修饰包括类似于常染色体水平的 H3 和 H4 乙酰化和 H3K4 甲基化（Kelly et al. 2002）。这表明，对于那些在两种性别的早期生殖细胞阶段都发挥功能的基因，存在尤为强大的选择压力以使其不定位于 X 染色体上。另外，已发现许多 X/常染色体类似基因倍增事件，其中仅有常染色体拷贝为生殖细胞所必需。在这些事例中，X 连锁拷贝在体细胞世系中具有功能，而在生殖细胞则非必需，因为在生殖细胞中有其常染色体拷贝的表达（Maciejowski et al.

2005）。总的来说，上述发现表明 X 染色体不适合在两种性别中都在早期生殖细胞发挥功能的基因的存在。只在卵子发生后期时所需要的基因（如只在雌性生殖细胞晚期所需要的；图 15-8）则能够逃脱雄性生殖细胞不联会产生的选择压力，因此它们能够被连锁于 X 染色体上。

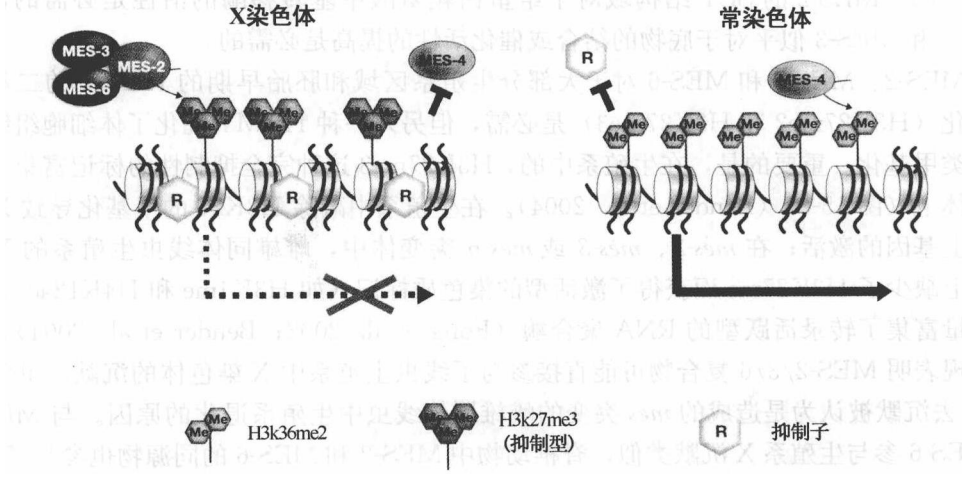



图 15-8　MES-2/3/6 和 MES-4 参与生殖系 X 沉默的模型

MES-2/3/6 复合物在 X 染色体上富集了抑制型染色质标记 H3K27me3，将 MES-4 排除于 X 染色体之外。MES-4 在常染色体上富集了一个不同的染色质标记（H3K36me2），可能抑制了一个未知的抑制子（黄色六边形），从而有助于这个抑制子在 X 染色体上的作用。这个结果表明 X 染色体的沉默是黄色六边形抑制子和 MES-2/3/6 催化的抑制型组蛋白甲基化的联合作用。

秀丽隐杆线虫中具有早期生殖系功能的基因显然被排除于 X 染色体之外，或者在 X 染色体上消失，或者经历基因复制后对常染色体拷贝中产生依赖。后者表明了一种可能的 X 染色体特异的遗传图谱的进化机制。在其他物种包括果蝇和哺乳动物中似乎有相似的压力作用于 X 连锁基因（Wu and Xu 2003）。但是现在我们还没有理解使生殖系必需基因离开 X 染色体的机制。

## 10. MES 组蛋白修饰酶对 X 染色体沉默的调控

前一节叙述了异染色质形成所产生的抑制是 XO 雄虫中单一 X 染色体所特有的，且由于 MSUD 效应，仅发生于减数分裂的粗线期。XX 雌雄同体线虫的 X 染色体在生殖系发育的其他阶段如何维持这种抑制型的染色质状态呢？针对母本效应不育（*mes*）突变体的遗传学筛选确定了 4 个参与 XX 线虫 X 染色体沉默的 *mes* 基因，同时它们可能也参与了 XO 雄虫中 X 染色体的沉默。遗传学和分子学研究结果表明编码的 MES 蛋白通过调节生殖细胞中 X 染色体上的部分蛋白修饰图谱而影响了 X 染色体的沉默。它们的功能对于生殖细胞的生存和发育是必需的。

三种 MES 蛋白：MES-2、MES-3 和 MES-6 形成一个类似于果蝇和脊椎动物中的 Polycomb 抑制型复合物（PRC2）的复合体共同发挥作用（参见 Xu et al. 2001；Bend-

er et al. 2004；Ketel et al. 2005；第 11 章）。MES-2 和 MES-6 是两个 PRC2 亚基，E(Z)（enhancer of zeste）和 ESC（extra sex combs）在线虫中的同源物（表 15-1）。E(Z) 和 ESC 与至少两个其他的伴侣蛋白一起催化了抑制型的组蛋白 H3K27 的甲基化。同样的，MES-2 和 MES-6 与一个新蛋白 MES-3 形成复合物并介导了 H3K27 的甲基化（图 15-8）。MES-2 的 SET 结构域对于组蛋白赖氨酸甲基转移酶的活性是必需的，而 MES-6 和 MES-3 似乎对于底物的结合或催化活性的提高是必需的。

MES-2、MES-3 和 MES-6 对于大部分生殖系区域和胚胎早期的 H3K27 的二和三甲基化（H3K27me2 和 H3K27me3）是必需，但另外一种 HKMT 催化了体细胞组织中的这类甲基化。重要的是，在生殖系中的，H3K27me3 这种完全抑制性的标记富集于 X 染色体上（图 15-8）（Bender et al. 2004）。在生殖系中敲除 H3K27 的甲基化导致 X 染色体上基因的激活；在 *mes-2*、*mes-3* 或 *mes-6* 突变体中，雌雄同体线虫生殖系的 X 染色体上缺少了 H3K27me 而获得了激活型的染色质标记（如 H3K4me 和 H4K12ac），并且大量富集了转录活跃型的 RNA 聚合酶（Fong et al. 2002；Bender et al. 2004）。这些发现表明 MES-2/3/6 复合物可能直接参与了线虫生殖系中 X 染色体的沉默。事实上 Xs 失去沉默被认为是造成的 *mes* 突变的雌雄同体线虫中生殖系退化的原因。与 MES-2 和 MES-6 参与生殖系 X 沉默类似，脊椎动物中 MES-2 和 MES-6 的同源物也参与了 XX 哺乳动物体细胞中的 X 失活（见第 11 和第 17 章）。

MES-2/3/6 复合物是如何被定位到需要被抑制的 DNA 区域，H3K27 甲基化又是如何抑制基因表达的呢？在果蝇中对于这些问题的解答已经浮现（见第 11 章）。在果蝇中，E(Z)/ESC 复合物被 DNA 结合蛋白 PHO 募集到一些被称为 polycomb 应答元件（PRE）的特殊位点上（Wang et al. 2004）。在复合物局部甲基化 H3K27 之后，基因抑制至少部分是由 H3K27me 处 polycomb 抑制复合物 1（PRC1）的募集所介导的（Wang et al. 2004）。奇怪的是，秀丽隐杆线虫的基因组并不包括 *pho* 或是大部分 PRC1 编码基因的明显同源物。因此，MES-2/3/6 复合物的定位以及它介导的基因抑制可能运用了与果蝇不同的机制。

第四个参与生殖系 X 抑制的 MES 蛋白是 MES-4。作为一种染色体结合蛋白，它有着全新的分布模式，且与预期完全相反。与其他 MES 蛋白不同的是，MES-4 以一种带状形式存在于五条常染色体，明显远离 X 染色体的大部分区域（图 15-8）（Fong et al. 2004）。像 MES-2 一样，MES-4 包含一个具有 HKMT 活性的 SET 功能域（Bender et al. 2006）。它对于生殖系和早期胚胎时期检测到的大部分的 H3K36 二甲基化（H3K36me2）是必需的。正如 MES-4 与常染色体的结合所预计的那样，H3K36me2 主要集中在常染色体上。在体细胞组织和后期胚胎中，另一个 HKMT 对于 H3K36 的甲基化是必需的。该活性似乎与转录延伸相关，这与酵母相一致。MES-4 介导的常染色体 H3K36 甲基化不依赖于转录，可能发挥了一种不同的，生殖系必需的表观遗传学作用（Bender et al. 2006）。

关键的问题是搞清楚 MES-4 是如何定位到常染色体上的，以及它在哪儿发挥什么作用。有趣的是，MES-4 被排除于 X 染色体之外需要 MES-2、MES-3 和 MES-6（图 15-8）；缺少其中任何一种 MES 蛋白，MES-4 将蔓延至 X 染色体上（Fong et al.

2002)。常染色体上 MES-4 的正常水平被认为参与了 X 染色体上的基因沉默。一个模型是 MES-4 或它的甲基化标记（H3K36me2）将一全基因组染色质抑制子（图 15-8 的黄色六边形）排除于常染色体之外，并通过这种方式将抑制作用聚焦于 X 染色体上(Fong et al. 2002)。MES-4 的缺失使这些抑制子离开 X 染色体结合到常染色体区域，导致 X 染色体的去沉默。MES2/3/6 复合物是 MES-4 排除的抑制子的合理候选物。然而 MES-2/3/6 催化的 H3K27 的甲基化分布在 *mes-4* 突变体中没有明显的变化。因此，关于其他可能参与这一过程的抑制子的研究仍在进行中。这个关于 MES-4 的作用模型类似于对酿酒酵母 Dot1 所提出的“通过抑制混乱而获得特异性”的模型（van Leeuwen and Gottschling 2002)。研究表明，Dot-1 介导的染色体 H3K79 甲基化能够将 Sir 抑制子排除，而使其作用聚焦于端粒。Dot-1 缺失使 Sir 从端粒处蔓延开，从而导致端粒的去沉默。

*mes* 表型是一个典型的与组蛋白修饰直接相关的表观遗传学表型。母本效应突变定义为突变表型不在第一代纯合突变体中显现，而出现在它们的后代中。在第一代 *mes/mes* 突变体中，*mes*/＋母源产生的并被装入卵母细胞中的一些野生型 MES 基因产物的存在足以使两个原始生殖细胞正确地扩增为一千多个具有功能的生殖细胞。然而这些具有生育能力的 *mes/mes* 线虫的生殖细胞不能为后代产生功能性的 MES 产物。结果在这些后代中，原始生殖细胞罕有增殖而发生退化。由 *mes-2* 和 *mes-4* 突变体编码的HKMT 活性必须建立起一个可遗传的，在两个最初原始生殖细胞的众多后代中完全维持的染色质状态。现有的 MES 蛋白功能的模型（图 15-8）是 MES-2、MES-3 和 MES-6 形成一个复合物，主要负责生殖细胞中 X 染色体上抑制型染色质修饰（H3K27me），直接参与 X 染色体的沉默。它们的作用是将 MES-4 排除于 X 染色体之外。因此，MES-4 包被了常染色体并对其进行 H3K36me 修饰，这被认为排除抑制活性并有助于将其聚焦于 X 染色体上。一般认为 MES 系统在母系生殖细胞和早期胚胎中发挥了表观遗传学作用，它建立正确标记的染色质区域，从而决定在幼虫生殖系发育过程中进行表达（常染色体区域）或沉默（X 染色体）（图 15-6)。MES 系统的缺失导致生殖系的死亡和不育，可能至少部分是因为 X 染色体的去沉默。

## 11. 精细胞 X 染色体的印记化

来自于不同配子的基因组汇入了一个受精卵中，从而使后者携带了大量不同的表观遗传学信息（图 15-6)。虽然 X 染色体在两种性别线虫的早期生殖细胞阶段是抑制型的，但 X 染色体在卵子发生过程中变成转录活跃的。相反，X 染色体也许能够说明为什么单亲遗传不会对线虫形成明显的损害。除了极小量的精子和 X 染色体上的内源生殖系表达位点，编码早期合子转寻物和那些对早期胚胎发育必须的基因在 X 染色体上都是严重不足的（Piano et al. 2000；Baugh et al. 2003)。因此，在 Xp 的抑制状态时，大部分编码较早期阶段必须蛋白的基因都不太可能出现在 X 染色体上，因此产生了遗传测试时不可预期的 X 染色体特异的单亲遗传。

然而 Xp 的抑制说明了早期胚胎中发生的不完全体细胞剂量补偿激活的原因。在 XX

胚胎的X染色体上，DCC成分的组装直到大约30细胞阶段才能被抗体检测到。这个现象稍滞后于Xp被H3修饰并完全富集，因此理论上是完全激活的。体细胞剂量补偿效应的激活应答于X连锁产物的水平，这种被称为X信号元件（XSE）的产物组成了X∶A比例的X部分。Xp上这些元件的全部或部分抑制可能导致功能性的XO早期胚胎（如X∶A比例为0.5）的产生。Xp在几轮细胞分裂之后重新被激活，X转录水平最终达到了一个2X染色体剂量的临界阈值，而X∶A比例为1，这样就产生了剂量补偿的级联效应。我们可以认为这是一个从母源/父源剂量补偿调控到受精卵调控的转换。

## 12. 总　结

至此，我们对秀丽隐杆线虫X染色体连锁位点的特异调节进行了详细描述。体细胞中该染色体特异的剂量调节被假定是由于不同性别中X倍性的不同而产生的，正如其他物种中发生的一样。有趣的是，线虫的生殖系和体细胞运用了不同的机制来应对X倍性的不同，这表明这些组织需要不同的X基因表达水平。在体细胞世系中，一个通常在有丝分裂和减数分裂过程中染色质参与致密化和分离过程的保守复合物似乎被用来特异性的一半抑制XX线虫中的两条X染色体。虽然以一个低效致密化复合物来实现转录效率降低一半在概念上非常诱人，但是为什么抑制作用局限于一半目前仍是一个谜。

生殖系中，在XX线虫生殖细胞发育的早期阶段和XO雄虫生殖细胞发育的所有阶段X染色体整条被沉默。染色质修饰物的MES系统参与了XX线虫的X沉默，同时也可能参与了XO雄虫中的X染色体沉默。这种沉默对线虫来说可以被承受，因为X染色体上没有多少生殖系表达基因。X连锁基因之所以极少对两种性别线虫的减数分裂都必需，被认为是由于X染色体在雄性减数分裂过程中缺少配对对象，而抑制性基因组防御机制又特异性针对不配对的染色体区域。生殖系中整条X染色体沉默的一个后果是不同性别的X染色体的倍性差异对这个组织并不产生什么影响。如果一些在两种性别中都起作用于的X连锁基因逃离沉默就会产生一个问题：XX和XO线虫的生殖系如何平衡它们的表达，对于这个问题的解答是未知的。生殖系中的平衡机制可能与体细胞组织中发挥作用的DCC不同。

秀丽隐杆线虫的生殖系和体细胞组织在许多基本方式上都不一样，本章强调了它们对X染色体调控的差异。一个有趣的论题是联合采用不同的已有机制，如在体细胞组织中利用致密化相关复合物微调X染色体表达而在生殖系中利用PRC2相关复合物沉默X染色体。雄虫中单一X染色体的异染色质化被认为显著改变了X染色体上的基因分布，以至于那些在生殖系发挥普遍功能和早期胚胎发育所必须的基因都几乎不出现在这条染色体上。因此，秀丽隐杆线虫的X染色体和它的调控代表了一个有趣的表观遗传学的染色体范围的调节和它对基因组进化的影响以及应答的交叉点。

吴　慧　译
张　慧　校
朱　冰　审校

# 参 考 文 献


Alekseyenko A.A. and Kuroda M.I. 2004. Filling gaps in genome organization. *Science* **303:** 1148–1149.

Baarends W.M., Wassenaar E., van der Laan R., Hoogerbrugge J., Sleddens-Linkels E., Hoeijmakers J.H., de Boer P., and Grootegoed J.A. 2005. Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. *Mol. Cell. Biol.* **25:** 1041–1053.

Baugh L.R., Hill A.A., Slonim D.K., Brown E.L., and Hunter C.P. 2003. Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome. *Development* **130:** 889–900.

Bean C.J., Schaner C.E., and Kelly W.G. 2004. Meiotic pairing and imprinted X chromatin assembly in *Caenorhabditis elegans*. *Nat. Genet.* **36:** 100–105.

Bender L.B., Cao R., Zhang Y., and Strome S. 2004. The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in *C. elegans*. *Curr. Biol.* **14:** 1639–1643.

Bender L.B., Suth J., Carroll C.R., Fong Y., Fingerman I.M., Cao R., Zhang Y., Briggs S.D., Reinke V., and Strome S. 2006. MES-4, an autosome-associated histone methyltransferase that participates in silencing the X chromosomes in the *C. elegans* germ line. *Development*. (In press.)

Carmi I., Kopczynski J.B., and Meyer B.J. 1998. The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex. *Nature* **396:** 168–173.

Chu D.S., Dawes H.E., Lieb J.D., Chan R.C., Kuo A.F., and Meyer B.J. 2002. A molecular link between gene-specific and chromosome-wide transcriptional repression. *Genes Dev.* **16:** 796–805.

Cowell I.G., Aucott R., Mahadevaiah S.K., Burgoyne P.S., Huskisson N., Bongiorni S., Prantera G., Fanti L., Pimpinelli S., Wu R., et al. 2002. Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. *Chromosoma* **111:** 22–36.

Csankovszki G., McDonel P., and Meyer B.J. 2004. Recruitment and spreading of the *C. elegans* dosage compensation complex along X chromosomes. *Science* **303:** 1182–1185.

Dawes H.E., Berlin D.S., Lapidus D.M., Nusbaum C., Davis T.L., and Meyer B.J. 1999. Dosage compensation proteins targeted to X chromosomes by a determinant of hermaphrodite fate. *Science* **284:** 1800–1804.

Fong Y., Bender L., Wang W., and Strome S. 2002. Regulation of the different chromatin states of autosomes and X chromosomes in the germline of *C. elegans*. *Science* **296:** 2235–2238.

Goday C. and Esteban M.R. 2001. Chromosome elimination in sciarid flies. *BioEssays* **23:** 242–250.

Goldstein P. and Slaton D.E. 1982. The synaptonemal complexes of *Caenorhabditis elegans*: Comparison of wild-type and mutant strains and pachytene karyotype analysis of wild-type. *Chromosoma* **84:** 585–597.

Haack H. and Hodgkin J. 1991. Tests for parental imprinting in the nematode *Caenorhabditis elegans*. *Mol. Gen. Genet.* **228:** 482–485.

Handel M.A. 2004. The XY body: a specialized meiotic chromatin domain. *Exp. Cell Res.* **296:** 57–63.

Hansen D., Hubbard E.J.A., and Schedl T. 2004. Multi-pathway control of the proliferation versus meiotic development decision in the *Caenorhabditis elegans* germline. *Dev. Biol.* **268:** 342–357.

Herrick G. and Seger J. 1999. Imprinting and paternal genome elimination in insects. *Results Probl. Cell Differ.* **25:** 41–71.

Hirano T. 2002. The ABCs of SMC proteins: Two-armed ATPases for chromosome condensation, cohesion, and repair. *Genes Dev.* **16:** 399–414.

Kelly W.G., Schaner C.E., Dernburg A.F., Ho-Lee M., Kim S.K., Villeneuve A.M., and Reinke V. 2002. X-chromosome silencing in the germline of *C. elegans*. *Development* **129:** 479–492.

Ketel C.S., Andersen E.F., Vargas M.L., Suh J., Strome S., and Simon J.A. 2005. Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. *Mol. Cell. Biol.* **25:** 6857–6868.

Maciejowski J., Ahn J.H., Cipriani P.G., Killian D.J., Chaudhary A.L., Lee J.I., Voutev R., Johnsen R.C., Baillie D.L., Gunsalus K.C., et al. 2005. Autosomal genes of autosomal/X-linked duplicated gene pairs and germ-line proliferation in *Caenorhabditis elegans*. *Genetics* **169:** 1997–2011.

Maine E.M., Hauth J., Ratliff T., Vought V.E., She X., and Kelly W.G. 2005. EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired DNA during *C. elegans* meiosis. *Curr. Biol.* **15:** 1972–1978.

Meyer B.J. 1997. Sex determination and X chromosome dosage compensation. In *C. elegans II* (ed. D.L. Riddle et al.), pp. 209–240. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

———. 2000. Sex in the worm counting and compensating X-chromosome dose. *Trends Genet.* **16:** 247–53.

———. 2005. X-chromosome dosage compensation. In *WormBook*, http://www.wormbook.org/chapters/www_dosagecomp/dosagecomp.html.

Nakayashiki H. 2005. RNA silencing in fungi: Mechanisms and applications. *FEBS Lett.* **579:** 5950–5957.

Nigon V. 1951. Polyploidie experimentale chez un nematode libre, *Rhaditis elegans maupas*. *Bull. Biol. Fr. Belg.* **85:** 187–255.

Piano F., Schetter A.J., Mangone M., Stein L., and Kemphues K.J. 2000. RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*. *Curr. Biol.* **10:** 1619–1622.

Powell J.R., Jow M.M., and Meyer B.J. 2005. The T-box transcription factor SEA-1 is an autosomal element of the X:A signal that determines *C. elegans* sex. *Dev. Cell* **9:** 339–349.

Prahlad V., Pilgrim D., and Goodwin E.B. 2003. Roles for mating and environment in *C. elegans* sex determination. *Science* **302:** 1046–1049.

Reinke V., Gil I.S., Ward S., and Kazmer K. 2004. Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. *Development* **131:** 311–323.

Reinke V., Smith H.E., Nance J., Wang J., Van Doren C., Begley R., Jones S.J.M., Davis E.B., Scherer S., Ward S., and Kim S.K. 2000. A global profile of germ line gene expression in *C. elegans*. *Mol. Cell* **6:** 605–616.

Schedl T. 1997. Developmental genetics of the germ line. In C. elegans *II* (ed. D.L. Riddle et al.), pp. 241–269. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Shiu P.K., Raju N.B., Zickler D., and Metzenberg R.L. 2001. Meiotic silencing by unpaired DNA. *Cell* **107:** 905–916.

Skipper M., Milne C.A., and Hodgkin J. 1999. Genetic and molecular analysis of *fox-1*, a numerator element involved in *Caenorhabditis elegans* primary sex determination. *Genetics* **151:** 617–631.

Turner J.M. 2005. Sex chromosomes make their mark. *Chromosoma* **114:** 300–306.

Turner J.M., Mahadevaiah S.K., Fernandez-Capetillo O., Nussenzweig A., Xu X., Deng C.X., and Burgoyne P.S. 2005. Silencing of unsynapsed meiotic chromosomes in the mouse. *Nat. Genet.* **37:** 41–47.

van Leeuwen F. and Gottschling D.E. 2002. Genome-wide histone modifications: gaining specificity by preventing promiscuity. *Curr. Opin. Cell Biol.* **14:** 756–762.

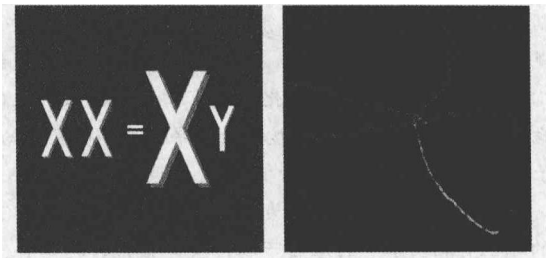
Wang L., Brown J.L., Cao R., Zhang Y., Kassis J.A., and Jones R.S. 2004. Hierarchical recruitment of polycomb group silencing complexes. *Mol. Cell* **14:** 637–646.

Wu C.I. and Xu Y. 2003. Sexual antagonism and X inactivation—The SAXI hypothesis. *Trends Genet.* **19:** 243–247.

Xu L., Fong Y., and Strome S. 2001. The *Caenorhabditis elegans* maternal-effect sterile proteins, MES-2, MES- 3, and MES-6, are associated in a complex in embryos. *Proc. Natl. Acad. Sci.* **98:** 5061–5066.

Yonker S.A., and Meyer B.J. 2003. Recruitment of *C. elegans* dosage compensation proteins for gene-specific versus chromosome-wide repression. *Development* **130:** 6519–6532.

# 第 16 章




# 果蝇中的剂量补偿效应

John C. Lucchesi[1] and Mitzi I. Kuroda[2]

[1]*Department of Biology, Graduate Division of Biological and Biomedical Sciences, Emory University, Atlanta, Georgia 30322*

[2]*Harvard-Partners Center for Genetics & Genomics, Brigham & Women's Hospital, Department of Genetics, Harvard Medical School, Boston, Massachusetts 02446*

## 本章目录

1. 剂量补偿现象在果蝇中的发现
2. 剂量补偿效应包括染色体修饰
3. 剂量补偿效应的调控源于 X 染色体和常染色体比例的测定
4. 非编码 roX RNA 促进 MSL 复合物在 X 染色体上的组装和定位
5. 染色体重塑活性的拮抗平衡
6. 展望

参考文献

## 概要

遗传信息转录的机制（如mRNA的合成）已得到了广泛的研究。一些能激活、维持和关闭基因状态的因子已经被发现。随着这些因子的出现，它们与调控序列和基因编码区的相互作用一直是研究的热点。在最近几年里，长期认为细胞分化和发育需要不同基因在时空上的协调调控的定律，已经使人们开始寻找一些基因活性相关的调控信号。早在这些研究开始之前，在果蝇中已经描述了一个协调调控的例子。在这个特殊的例子中，这些受协调调控的基因并非与功能相关，而是与遗传物质有相同的定位：它们都出现在一条X性染色体上。这种调控是为了确保有两条X染色体的雌性和一条X染色体的雄性有相同水平的基因产物，换句话说，就是补偿性别间与X染色体连锁的基因量的差异。在研究这种调控水平时，问题已由“这些不相关的基因是如何协调调控的?”变为“调控整条染色体活性的机制是什么?”

果蝇中剂量补偿效应的研究——加强雄性单条X染色体中大部分基因转录的机制，揭示包括位点特异性的组蛋白乙酰化、X染色体特异的非编码RNA（称为roX1和roX2）和染色体范围内进化保守的染色质修饰过程的靶点，该靶点被显示为如标题图示中在X染色体上的定位。

## 1. 剂量补偿现象在果蝇中的发现

多数生物体的核型（即全部染色体）中都有一对性染色体。在果蝇中，雌蝇有两条称为 X 染色体的性染色体，这两条染色体的形状和遗传信息已经被识别，它们都在所有体细胞中被激活。雄蝇有一条 X 染色体和一条 Y 染色体，Y 染色体在形态学和遗传信息上都与 X 染色体不同（图 16-1）。在性染色体上，有一些负责性别决定和性别分化的基因，X 染色体上还有一些参与基本的细胞功能或发育通路的基因。含两条 X 染色体的雌性有两倍量的这些基因，而只有一条 X 染色体的雄性中只有一份。但在两性中，大部分这些基因的产物水平是相同的。早在 1903 年，H. J. Muller 在研究 X 连锁基因部分缺失突变个体眼睛色素水平时就在果蝇中首次发现了这个矛盾（图 16-2）。Muller 认为一定存在一种调控机制来补偿果蝇中 X 染色体连锁基因在雄性和雌性中差异，使 X 染色体连锁基因在两性中的水平相同。他把这种假定的调控机制称为“剂量补偿效应（dosage compensation）”。

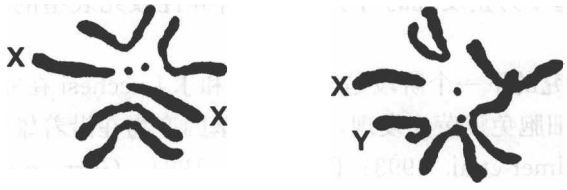



图 16-1　由显微描绘器绘制的雄性（右边）和雌性（左边）果蝇分裂中期相 X 染色体占核型的 20%（经许可转印自 Morgan 1932）。

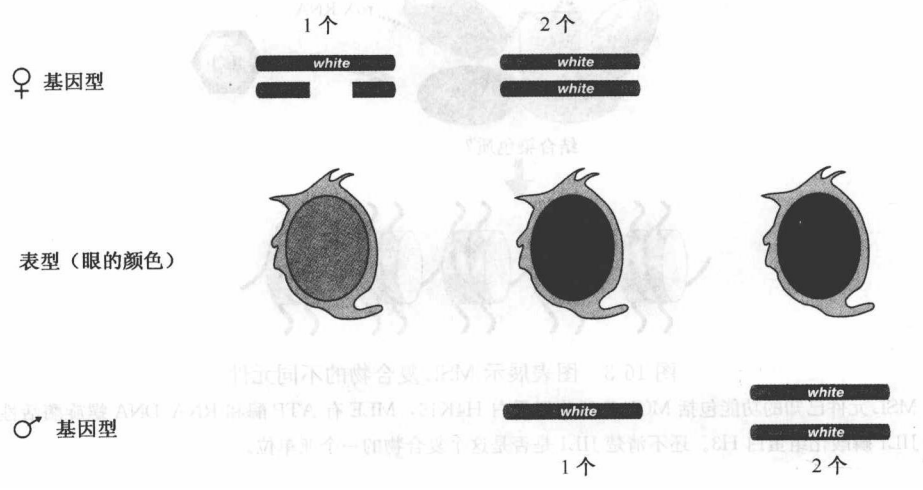



图 16-2　Led H. L Muller 描述剂量补偿效应假说的图示

X 染色体连锁基因 *white* 的突变的等位基因（$w^a$）是一个亚效等位基因，负责眼睛局部色素的合成，它出现在 X 染色体上用深色的框表示。色素形成的水平与每种性别中 $w^a$ 的剂量成正比，但由于剂量补偿效应，有一份 $w^a$ 剂量的雄性和两份 $w^a$ 剂量的雌性中色素的数量相当。

30 多年后，A. S. Mukherjee 和 W. Beermann 通过调控 X 染色体连锁基因的转录首次证明了剂量补偿效应。幼虫唾腺巨型多线染色体转录的放射自显影（autoradiography）技术是一种能呈现转录实时情况的分子技术，利用这种技术可观察到掺入雄蝇单条 X 染色体中的 [$^3$H] 尿嘧啶与掺入到雌蝇两条 X 染色体的数量是相同的。因此看起来，在雄蝇单条 X 染色体上 RNA 合成量大约是雌蝇每条 X 染色体合成量的两倍。另一个实验的突破是 J. Belote 和 J. Lucchesi 识别出了四种基因，这几种基因在雌性中的缺失不受影响，但在雄性中的缺失却是致死的，雄性 X 染色体上出现大约正常水平一半的 [$^3$H] 尿嘧啶掺入量。另外，X 染色体失去它正常的界限，并呈现疏松的外观（puffed），这被解释为是雌性各条 X 染色体转录水平增加的象征。结果说明，X 染色体连锁基因产物等量是由于雄性 X 染色体的转录活性是雌性每条 X 染色体转录活性的两倍来完成的。这四种基因中，两种是新发现的（雄性特异性致死基因 1，*male-specific lethal*，*msl1*；雄性特异性致死基因 2，*msl2*），另两种基因（无雄基因，*maleless*，*mle*；雄性特异性致死基因 3，*msl3*）是之前在对自然群体的其他研究中发现的（在 1987 年 Lucchesi 和 Manning 的研究中可以找到早期研究剂量补偿效应的专门文献）。为便于参考，迄今为止发现的可引起雄性特异性致死表型的所有缺失突变基因产物都叫做 MSL。

剂量补偿作用研究的下一个阶段是 B. Baker 和 J. Lucchesi 在完成三种 *msl* 基因后，开始研究 *mle*。通过细胞免疫荧光发现，这四种基因产物在沿着雄性多线 X 染色体的很多位点上共定位（Palmer et al. 1993；Bone et al. 1994；Gorman et al. 1995；Kelley 等 1995；Gu et al. 1998）。观察结果和不同基因产物对 X 染色体结合的相互依赖说明它们形成一个复合物（图 16-3）。

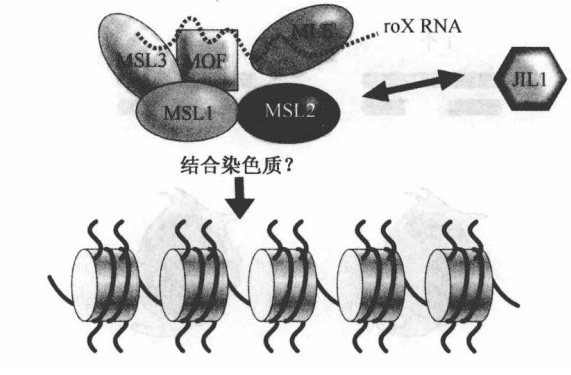



图 16-3 图表展示 MSL 复合物的不同元件

MSL 元件已知的功能包括 MOF 乙酰化组蛋白 H4K16，MLE 有 ATP 酶和 RNA/DNA 螺旋酶活性，JIL1 磷酸化组蛋白 H3。还不清楚 JIL1 是否是这个复合物的一个亚单位。

## 2. 剂量补偿效应包括染色体修饰

雄性 X 染色体上 MSL 复合物的出现与组蛋白 H4 Lys16 特异的乙酰化异构形式有

关（Turner et al. 1992；Bone et al. 1994）。酵母中，组蛋白 H4 特定的共价修饰已经证明对沉默染色质和激活染色质之间状态的维持起重要作用。负责 H4K16ac 的组蛋白乙酰化酶（histone acetyltransferase，HAT ）Sas2 功能的丧失，会导致端粒异染色质蔓延到邻近亚端粒染色质上（Suka et al. 2002；见第 4 章）。果蝇中发现 H4K16ac 在雄性中沿着 X 染色体高度富集。最近结构研究已经表明，核小体间相互作用可发生在一个核小体上组蛋白 H2A-H2B 二聚物的酸性的片段上，以及从邻近的核小体延伸而来的带正电的组蛋白 H4 尾巴片段（16～26 个残基）上（Schalch et al. 2005）。当 Lys16 被特异的乙酰化后，所带正电荷被中和，这就减弱了核小体间的结构的抑制，从而影响剂量补偿效应。

最初发现的参与剂量补偿效应的因子中没有具乙酰化酶作用的，因为早期的突变工作主要集中于常染色体上寻找。从寻找雄性特异性致死的 X 连锁基因开始，才识别出一个新基因 *mof*(*males absent on the first*)（Hilfiker et al. 1997）。*Mof* 编码的重组蛋白是一种乙酰化酶，能特异地乙酰化组蛋白 H4K16(Akhtar and Becker 2000)。它是 MSL 复合物的一部分，仅作用于雄性 X 染色体上的 H4(Smith et al. 2000)。MOF 是乙酰化酶 MYST 家族中的一类亚家族分支。这个亚家族都具有 chromo 结构域，可被分为 H4K16 体内特异性乙酰化酶和像酵母中 Esa1 那样能乙酰化 H4 末端四个赖氨酸的酶(Smith et al. 1998)。另一种 MYST 家族的成员是 SAS2，在酵母中可特异的乙酰化 H4K16，但无 chromo 结构域。

在 MOF 被发现和研究的同时，人们证明了一个具有潜在酶活作用的其他亚单位 MLE。MLE 具有 RNA/DNA 螺旋酶、三磷酸腺苷酶（ATPase）和体外单链 RNA/ssDNA结合活性（Lee et al. 1997），这暗示了 RNA 可能在 MSL 中的功能。一些多蛋白复合物可与染色质作用，通过 ATP 水解作用改变核小体的构象，控制常染色质基因的转录效率。ATP 水解作用是含有螺旋酶结构域的蛋白家族的功能（Sif 2004），尽管它在 MSL 复合物中不一定也有相同的功能，MLE 是这个家族的代表。MLE 的直向同源物（ortholog）包括人类的 RNA 螺旋酶（RNA helicase A，RHA），它属于 DEXH RNA 螺旋酶亚家族，其特征是包含新的结构域（Kuroda et al. 1991）和 N 端一段约 350 个氨基酸的保守结构（Pannuti and Lucchesi 2000）。除了在染色质重塑中的作用外，MLE 还可能通过改变 RNA 的结构在剂量补偿机制中起作用。

现在仍没有完全了解 MSL 复合物其他部分的特殊功能。MSL3 是 MSL 复合物中的一个亚单位，具有非经典的 chromo 结构域和 chromoshadow 结构域（Aasland and Stewart 1995；Koonin et al. 1995），是与含有 chromo 结构域的 HAT 协同进化的一类蛋白家族中的一员（Pannuti and Lucchesi 2000）。酵母中的 Eaf3(Esa1 相关蛋白因子 3)是这个家族中的一员，它与 MYST 酶 Esa1 共同在 NuA4 复合体中（Eisen et al. 2001）。在果蝇中，Tip60 多蛋白复合物包括 MSL3 的旁系同源物（paralog）MRG15(Kusch et al. 2004)。人类 MAF2 复合物中 MRG15 与 MOF 相关（Pardo et al. 2002）。更有趣的是，人类中存在与 MSL 复合物相关的其他复合物，包括 MOF、MSL1、MSL2 和 MSL3 的人同源物。在人类细胞中，这个复合物能特异性乙酰化组蛋白 H4K16，并负责多数组蛋白的乙酰化修饰，可包括两个不同基因编码的果蝇 MSL3 同

源物的三种不同的形式之一（Smith et al. 2005）。

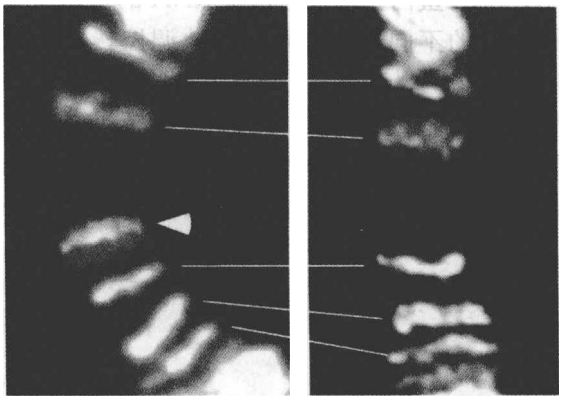
在果蝇中，MSL1 和 MSL2 介导 MSL 复合物连接到染色体上（Kelley et al. 1995；Copps et al. 1998；Gu et al. 1998）。没有发现这些蛋白中含有已知的 DNA 结合结构域，也没有发现它们在染色质上的协同作用。MSL1 通过 N 端的螺旋-螺旋结构域与 MSL2 的 RING 指状结构域相互作用（Copps et al. 1998；Scott et al. 2000），通过靠近 C 端的保守的结构域与 MSL3 和 MOF 相互作用（Morales et al. 2004）。MSL2 和 X 染色体的关系是非常稳定的（Straub et al. 2005c）。MSL1 和 MSL2 在人类中同源物已经被报道（Marn 2003）。因为果蝇中 MSL1 和 MSL2 负责 MSL 复合物和 X 染色体的联系，因此我们有理由推测 hMSL1 和 hMSL2 可能在将人类 MSL 复合物引导到其作用位点的过程中起一定作用。

除了雄性特异性的因子外，一些在两性中参与染色体组织和转录的一般因子也可能参与剂量补偿效应。串联激酶（tandem kinase）JIL-1 在雄性和雌性中沿着所有染色体分布，但它在雄性 X 染色体上分布更集中。这种富集依靠 MSL 复合物。JIL-1 介导组蛋白 H3 的磷酸化，并在基因组的转录激活区域维持染色体结构的打开状态（Jin et al. 1999；Wang et al. 2001）。但 JIL-1 在剂量补偿效应中起特异功能的还是一般功能仍未确定。现在已经发现包括核孔蛋白（nucleoporin）和外核体（exosome）复合物等亚单位可以与一些 MSL 发生免疫共沉淀（Mendjou et al. 2006），但这些蛋白在剂量补偿效应中的作用仍未确定。

复合物的活性必须进化到能控制一些先前存在的基因，每个复合物都有它们特异的调控信号和本来的表达水平。插入 X 染色体上的调控转基因仅诱导表达时才出现 MSL 的异位结合，这时 MSL 复合物才可能被激活的基因活化（图 16-4）（Sass et al. 2003）。总的转录是在哪步被上调两倍呢？在雄蝇中，H4K16ac 不仅仅局限在代偿 X 染色体连锁基因的启动子区域，也被发现沿着整个 X 染色体连锁基因转录单位分布，这些转录单位被 MSL 复合物激活，后者的乙酰化水平在启动子区域相对一般，在中间或 3′区较高（Smith et al. 2001）。这种乙酰化的分布说明 MSL 复合物通过促进转录延伸和 RNA 聚合酶可能的循环起效，而不是通过直接加强启动子的可及性（accessibility）来增加 X 染色体连锁基因的表达。

已经提出了一种候选的假说来解释 MSL 复合物的功能（Birchler et al. 2003）。这个假说的根据是雌性中所有染色体的活性是建立在 MOF 的均匀分布的基础上的。雄性中由于缺少一条 X 染色体，因此在常染色体和单条 X 染色体上的 MOF 和其他未知的因子就会有更高的浓度，这就使它们具有更高的转录活性，形成的 MSL 复合物被招募到 X 染色体上来使包括 MOF 在内的高转录的因子从常染色体上离去。在这个模型中，MSL 复合物的根本作用是下调常染色体上基因的表达，而不是上调 X 染色体上的连锁基因。但是，已有两个独立的研究与此假说相矛盾。在培养的雄性组织细胞中，缺失 MSL 复合物的靶基因可减少 X 染色体上基因的转录水平，但常染色体上基因的表达却不受影响（Hamada et al. 2005；Straub et al. 2005b）。最近，雄性生殖细胞系中 X 染色体连锁基因的表达也有类似的结果（Gupta et al. 2006）。

染色体组装可以被细分为依赖复制的核小体植入和不依赖复制的核小体植入。后者



活化的基因 (+GAL)　　失活的基因 (–GAL)

图 16-4　MSL 复合物靶向性到活化的基因

一个反式激活蛋白 GAL4 控制下的启动子结构已经插入到幼虫唾腺染色体中缺乏 MSL 复合物的位点（右图）。当引入一个表达 GAL4 的基因到基因组中时（左图），反式激活蛋白结合到这个结构（红颜色）上，并招募 MSL 复合物（绿颜色）（经许可转印自 Sass et al. 2003 [©：美国国家科学院]）。

发生在染色体的转录活性区，包括组蛋白 H3 被变体 H3.3 替换（参考第 13 章）。与这个现象相一致的是，在雄性细胞中，组蛋白 H3.3 加入到 X 染色体的程度比加入到常染色体中的程度高（Mito et al. 2005）。

## 3. 剂量补偿效应的调控源于 X 染色体和常染色体比例的测定

每个胚胎都需要计数其 X 染色体的数量以便做一个重要的决定：是否需要剂量补偿效应。如没有上调雄性单条 X 染色体，或没有下调雌性两条 X 染色体，这个不正确的决定可能致死。在果蝇中，计算 X 染色体的过程与性别决定的过程是相配合的（详情参考 Cline and Meyer 1996）。性别由每个细胞核中 X 染色体数量决定，如 XX 胚胎是雌性，XY 胚胎是雄性。雄性 Y 染色体是雄性生殖所需要的，不像在哺乳动物中，它对性别的表型没有作用。形式上，是 X 染色体和常染色体的比例控制着性别和剂量补偿效应，如计算 X 染色体的机制对常染色体组的数量很敏感一样。这种情况在 2X∶3A 三倍体中更加明显，如 X∶A 的比例在雄性 XY∶2A 和雌性 XX∶2A 之间。2X∶3A 三倍体的差别在于作为一个雌雄间体，它的雄性细胞和雌性是混在一起的。

X 与 A 的比例通过调控二元开关基因——性别致死基因（*Sex lethal*，*Sxl*），控制性别决定和剂量补偿效应。*Sxl* 编码一个雌性特异性的 RNA 结合蛋白，在性别决定和剂量补偿通路中分别调控着关键 RNA 的剪切和转录过程（图 16-5）。*Sxl* 是 X 染色体编码的，受 X 染色体编码的一些转录因子正调控，因此，有两条 X 染色体的胚胎可通过早期受调控的启动子 $P_e$ 起始表达 *Sxl*，但每个核只有一条 X 染色体的胚胎就不能从 $P_e$ 上起始表达。在早期胚胎 *Sxl* 激活过程中这个最初瞬间的不同，可以通过自我反馈

来稳定，在这个循环中，SXL 蛋白正调控它自身 RNA 的剪切而其启动子是组成型表达的。雌性模型中，*Sxl* 起始的不同通过性别特异地调控转换基因（*transformer*，*tra*）的剪切来实现。反过来，这些基因产物与其他基因产物，如转换基因 2(*tra2*) 一起出现在两性中，指导双重性别基因（*doublesex*，*dsx*）起始转录成一定量的调控基因，抑制雄性特异性基因的表达，实现雌性性别的分化。在雄性胚胎中，则是通过 *dsx* 错误转录剪切成为抑制雌性特异性基因的表达的产物，从而实现向雄性的分化。

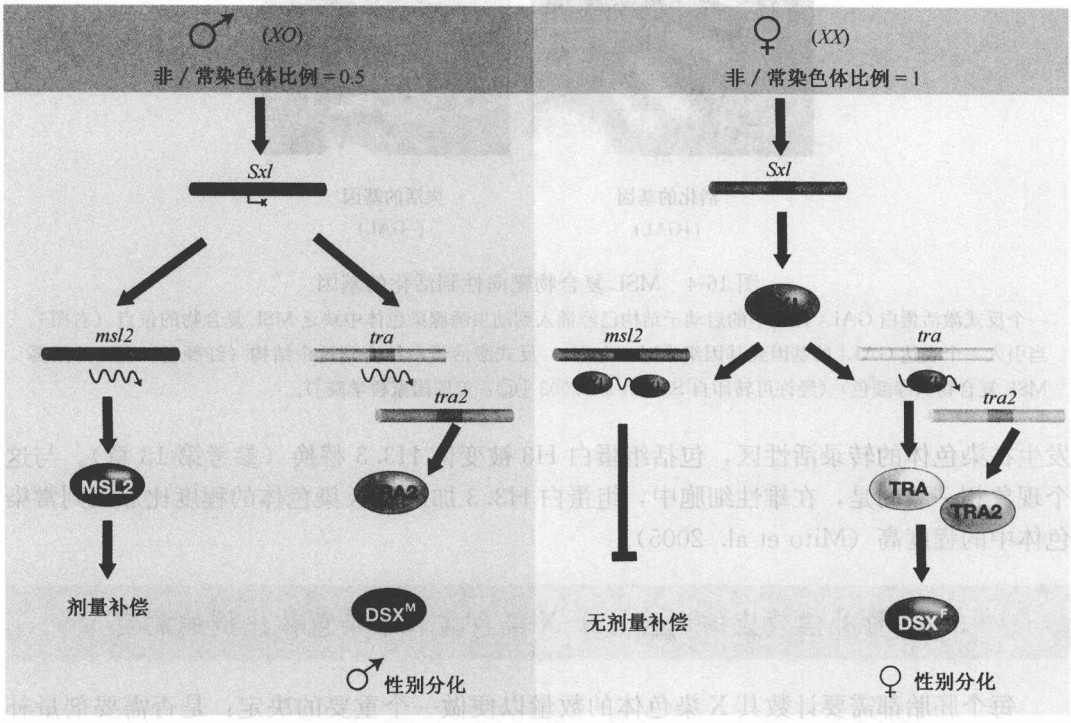



图 16-5 控制性别决定和剂量补偿效应的图示

如果 X/A 的比等于 1，级联调节作用使雌性性别发育。在雌性中，*Sxl* 基因产物的出现抑制 *msl2* 的翻译和 MSL 复合物的组装。如果 X/A 的比仅有 0.5，级联调节作用的缺乏，MSL 复合物形成雄性性别发育。

在剂量补偿通路中，SXL 的一个关键靶子是 *msl2* mRNA(Bashaw and Baker 1997；Kelley et al. 1997)。SXL 的结合位点在 *msl2* mRNA 的 5′UTR 和 3′UTR 上。SXL 通常仅在雌性中出现，通过与 *msl2* mRNA 5′UTR 的作用，抑制 *msl2* mRNA 的翻译（图 16-5）。如果 SXL 在雌性中缺失，剂量补偿效应异常开启，雌性将死亡。相反，如果在雄性中异常表达 SXL，就会关闭剂量补偿效应，导致雄性死亡。雌性中异位表达 MSL2 足以使 MSL 复合物在两条雌性染色体上组装，这也说明 MSL 复合物的其他组成成分或者开启或者不受 MSL2 表达的影响。MSL1 和 MSL3 在雌性中正常表达，但在缺失 MSL2 蛋白时它们的表达却不稳定。MLE 和 MOF 在雌性中稳定表达，当缺失 MSL2 时对 X 染色体变得没有特异性亲和力。

总之，剂量补偿效应必须是对核中 X 染色体数量的反应。在胚胎发育的早期，X 染色体的计数工作已经完成。雄性表达 MSL2 蛋白、缺乏 MSL 复合物，但雌性抑制

MSL2 的翻译，以防止两条 X 染色体出现不适当的剂量。

### 4. 非编码 roX RNA 促进 MSL 复合物在 X 染色体上的组装和定位

哺乳动物和果蝇中，剂量补偿效应最吸引人和最神秘的方面之一是非编码 RNA 能引导正确染色体上发生代偿（详情参考 Gilfillan et al. 2004；Kelly 2004；Straub et al. 2005a；第 17 章）。两种称为 roX(RNA on X) 的非编码 RNA，虽然长度和序列不同，却具有功能上的冗余性，即在果蝇中都能把 MSL 复合物引导到雄性 X 染色体上（Meller and Rattner 2002）。传统的突变筛选通常没能发现这些像 roX RNA 一样的编码产物有相同功能的基因的存在。但在成熟脑中 roX RNA 是作为一种雄性特异性的 RNA 而被偶然发现的（Amrein and Axel 1997；Meller et al. 1997）。进一步研究发现，这两种 RNA 没有明显的 ORF，也不能与 MSL 复合物沿着整条 X 染色体共定位。直到 X 染色体上 roX1 和 roX2 双突变时，roX RNA 的功能才被揭示。roX 双突变中大部分雄性致死，并伴有 MSL 复合物的严重异位，但单个突变却没有这种表型（Meller and Rattner 2002）。令人惊讶的是这两种 roX RNA 大小不同（3.7kb 和 0.5～1.4kb），并几乎没有相似的序列。过表达 MSL 蛋白可部分克服 roX RNA 的缺失，说明这种蛋白拥有剂量补偿效应中所有的关键功能，但它需要这些 RNA 来组装或定位（Oh et al. 2003）。

roX RNA 可通过与 MSL 蛋白免疫共沉淀获得，这说明两者的物理性质上的相关性（Meller et al. 2000；Smith et al. 2000）。现在仍不清楚 roX RNA 是共存的，还是形成两类不同的 MSL。roX1 的部分删除揭示单纯的缺失 300 个核苷酸不影响其功能，这说明其内部序列是重复的（Stuckenholz et al. 2003）。roX RNA 显示惊人的量的可变性，这说明它们可能不是在 MSL 复合物中形成一个不变的结构，而是装饰这个复合物的表面。部分纯化该复合物说明存在由 MSL1、MSL2、MSL3 和 MOF 蛋白组成的紧密核心，以在极低的盐浓度下才不保留的 roX RNA 和 MLE 螺旋酶（Smith et al. 2000）。在体外缺失 roX RNA 的最小蛋白核心复合物仍可特异性地乙酰化核小体 H4K16(Morales et al. 2004)，这与 roX RNA 参与组装或开启剂量补偿效应，而不是在基因调控中起直接作用的观点相一致。

MSL 复合物是怎样结合到 X 染色体上的，它又是如何识别的呢？缺少两种 roX RNA，或 MLE、MSL3、MOF 中的一个，可使部分 MSL 复合物结合于 35～70 个位点上（图 16-6）。在 *mle*、*msl*3 和 *mof* 突变中比较这些“高亲和力”的位点，发现它们在细胞学水平几乎相同，但仍不清楚它们在分子水平上的共性。*mle* 突变中重要的缺陷可能是不能把 roX RNA 加入到复合体上，因为在缺少 MSL3 或 MOF 时，这种 RNA 可与部分 MSL 复合物共定位，而缺少 MLE 时却不能（Meller et al. 2000）。在缺少 MSL1 或 MSL2 时，剩下的 MSL 蛋白或 roX RNA 没能保持对 X 染色体的识别能力，这导致一种假说：MSL1 和 MSL2 至少共同具有对 X 染色体靶向性的起始特异性。然而，这两种蛋白都没含有已知的 DNA 结合基序，在体外也没有证明它们与 DNA 的结合。

X 染色体上吸引这个复合体的特异性信号可能是什么呢？有两个可能的基本模型来

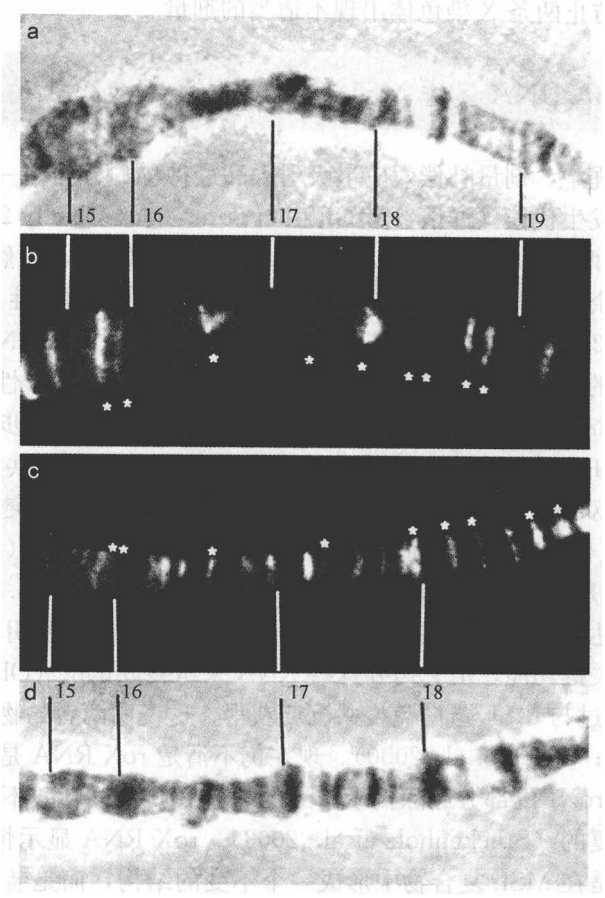



图 16-6 “高亲和力”的位点

来自 *msl3* 等位基因功能缺失幼虫（a，b）和野生型幼虫（c，d）的唾腺染色体。a 和 d 为相同相位的图像，b 和 c 表示对同一染色体上复合物的免疫染色。b 图中的这个复合物（绿色）位点的数量比 c 图少（经许可转印自 Demakova et al. 2003[© Springer]）。

调控整条染色体。①单个的位点或是很少的位点可能顺式控制这条染色体，就像是哺乳动物中的 X 染色体失活过程一样（参考第 17 章）。这种机制需要这个复合物在空间上被限制到核内一个特定的区域，或者通过一些因子从中心控制区延伸到染色体的其他区域进行远距离的调控。②染色体可以识别分散在它的全长中独特序列。这样，染色体的任何部分都可以被自主的调控。任何序列都可能是这两种模型的元件，包括具有远程调控能力的分散序列和中心控制区的指令。

roX RNA 通常是 X 染色体编码的，*roX1* 基因在 X 染色体的尖部附近，*roX2* 基因在 X 染色体上常染色质部分的中间区附近。像是哺乳动物中的 *Xist* 基因一样，位于 X 染色体上的 *roX* 基因可能是为了引导 MSL 复合物装配到该染色体上。当通过转基因把 *roX* 移到常染色体上，它也能吸引 MSL 蛋白到它的主要插入位点（图 16-7），在这里，这个复合物可顺式延伸至两侧的序列中（Kelley et al. 1999；Kageyama et al. 2001）。在特定的遗传条件下，如 X 染色体上没有竞争的内源 *roX* 基因，时常可以看到 *roX* 转基因在常染色体

上的广泛延伸（图 16-8）（Park et al. 2002）。这种广泛的延伸可通过过表达有限的 MSL 关键蛋白 MSL1 和 MSL2 而增加，也可以通过过表达竞争性转基因 roX RNA 来减少这种情况，这说明 MSL 复合物成功的共转录装配可以促进局部的延伸（Oh et al. 2003）。在核内有效的组装功能 MSL 复合物可能是 roX RNA 的主要功能。X 染色体上 *roX* 基因最初的组装可能提高 MSL 有效的定位到 X 染色体上的效率，但因为 roX RNA 可反式作用，因此它明显不是 MSL 最终定位所必需的（Meller and Rattner 2002）。

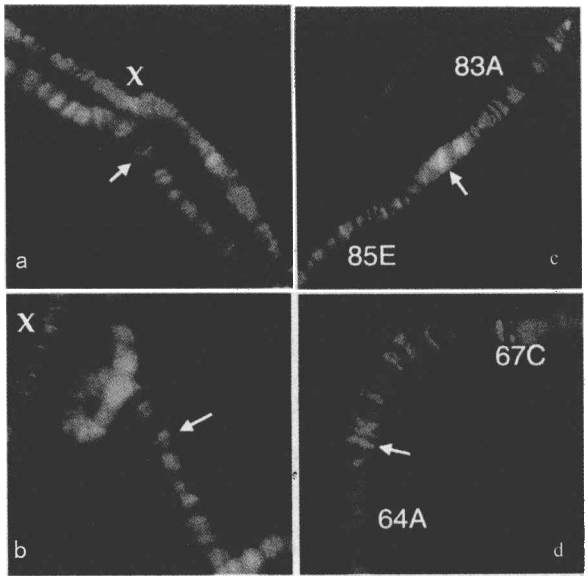



图 16-7　MSL 复合物从常染色体上转基因的延伸与内源性 X 染色体连锁基因 *roX* 功能的减少是成比例的

当转基因是 roX RNA 的唯一来源时，常染色体达到最大的延伸。(a) 雄性野生型的 X 染色体；常染色体转基因的出现用 MSL 带的箭头表示（红色）。(b) 仅有一个激活的 X 染色体连锁的 *roX* 基因的雄性。相比于野生型，MSL 复合物仅轻度延伸。(c，d) 在带两个不同的转基因和 X 染色体上两个 *roX* 基因都敲除的雄性中的延伸。

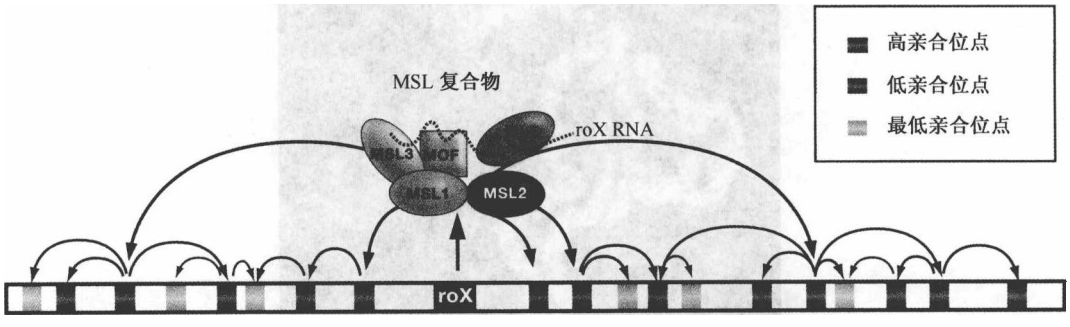



图 16-8　MSL 复合物靶向性到 X 染色体上的模型

复合物在 *roX* 基因转录的位点组装，然后与 X 染色体上亲和力或高或低的不同位点相结合。该模型也适合常染色体上该复合物从异位的 *roX* 基因上的延伸。

除了前面提到的高亲和力位点，该复合物在沿着 X 染色体的其他许多位点上都明显表现出了不同水平的亲和力（Demakova et al. 2003）。在稳定的易位果蝇中，还没有证明 MSL 复合物可从 X 染色体上延伸至邻近的常染色体序列中（Fagegaltier and Baker 2004）。因此，如果 MSL 复合物延伸是屏蔽 X 染色体的主要机制，那么 X 染色体可与 MSL 复合物的结合程度可能远远超过常染色体。事实上，通过检测 X：A 的异位和仅含一套 X 染色体的转基因实验，发现 X 染色体上大部分 30kb 的片段都有吸引 MSL 复合物的能力，这种作用在更小的片段中还未确定（Oh et al. 2004）。

怎样把不同的观察结果形成一个模型来解释 MSL 复合物靶向性到 X 染色体上？图 16-9 描述的 X 染色体识别模型是与已有数据最吻合的模型。在这个模型中，MSL 复合物在 roX RNA 转录的位点组装，随后根据 MSL 复合物与 X 染色体上不同位点的相对亲和力，到达转录位点两侧或更远的位点。仍不清楚的是，许多或大部分靶向性通常是以反式的方式单独发生，还是通过局部顺式延伸发生。但可以明确的是，与常染色质相比，MSL 复合物更易与 X 染色体结合。尽管现在没有发现 X 染色体上所特有的序列，但随着多种物种果蝇基因组的比较，可能将揭示一些简并的序列。

## 5. 染色体重塑活性的拮抗平衡

雄性多线 X 染色体对一些通用的非染色体特异性的染色质调控因子的剂量或活性的改变特别敏感。例如，*Iswi* 缺失突变揭示了 ISWI 的复合物和 MSL 复合物功能的相互作用。该突变导致了雄性 X 染色体上整个结构的改变：在唾腺标本中，染色体看起来特别的短粗（图 16-9），MSL 复合物和 H4K16ac 呈现明显连续的形式而非细带。（Deuring et al. 2000）。这种细胞学现象的分子基础仍不清楚，但与一个模型相符，在这个模型中，染色质特殊区域失去了组织性。正常的多线染色体有一个特征性的形式：相对浓缩的带、去浓缩的带间和疏松区，这种组织形式在 *Iswi* 突变的雄性果蝇 X 染色体上大部分丢失。X 染色体的变短可能是由于结构的丢失和浓缩导致宽度的异常增加。

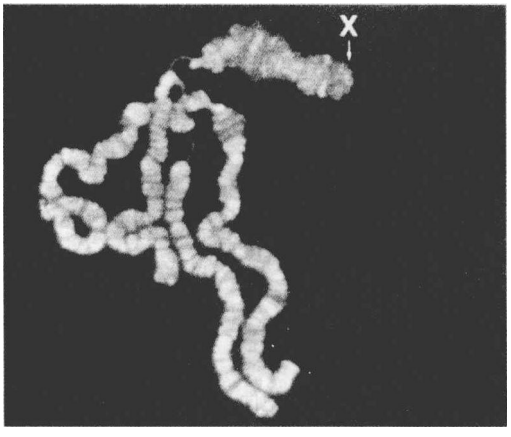



图 16-9 *Iswi* 基因的下效突变首先影响雄性 X 染色体的结构组织

在野生型雄性中，X 染色体通常看起来像常染色体的一条臂（经许可转印自 Deuring et al. 2000[© Elsevier]）。

模拟开关蛋白（imitation switch protein，ISWI）是一种 ATP 酶，存在于果蝇三种染色质重组复合物中：NURF(nucleosome remodeling factor，核小体重塑因子)，ACF(ATP-dependent chromatin assembly and remodeling factor，ATP 依赖的染色质组装和重塑因子）和 CHRAC(chromatin accessibility complex，染色质趋近性复合物）(Smith and Peterson 2005)。体外研究已经证明 ACF 和 CHRAC 使核小体有序排列，而 NURF 则打乱核小体排列。在体内 ACF 和 CHRAC 可作为染色质组装因子启动染色质整体的形成和抑制染色质特定的位点。相反，NURF 改变核小体体外重建时 hsp70 启动子处的位置，如果核小体被首先超乙酰化，其重塑活性也会大大加强。尽管 NURF 在标记染色质的可及性上具有明显的生物化学活性，但 *nurf301* 的突变也表现出 *Iswi* 突变中雄性 X 染色体异常去浓缩的现象（Badenhorst et al. 2002)。

在雌性转基因果蝇中，诱导 MSL 复合物出现的地方也能看到 *Iswi* 和 *nurf* 突变时唾腺 X 染色体的缺陷，这种缺陷在 MSL 复合物失活（如缺少 H4K16ac）的雄性果蝇中未被发现。另外，与纯化的 ISWI 体外竞争实验说明，组蛋白 H4 尾部 K16 的乙酰化可以减弱 MSL 复合物与分离的核小体的结合能力（Corona et al. 2002)。这些研究者推测可通过减少 ISWI 促进常规核小体有序形成的能力，使 MOF 起作用改变了 X 染色体连锁基因的表达（Corona et al. 2002)。

当异染色质蛋白 SU(VAR)3-7 过表达时，可以看到一个完全不同的染色体表型(Delattre et al. 2004)。SU(VAR)3-7 是异染色质的一个结构组成，它可与 HP1(heterochromatin protein 1，异染色质蛋白 1）和组蛋白赖氨酸甲基转移酶（histone lysine methyltransferase，HKMT) SU(VAR)3-9 共定位（参考第 5 章)。尽管过表达 SU(VAR)3-7 可以引起雄蝇和雌蝇染色体的形态学变化，但由于雄蝇 X 染色体形状更小并更致密，因此它受影响的程度更大。

令人好奇的是，雄性多线 X 染色体比其他染色体更敏感，在受到正负作用时更容易改变染色质组件的平衡。维持染色质状态继承的重要机制可能是在 DNA 合成后，对组装因子的竞争。由于雄性多线 X 染色体一直处于开放的状态，当主要染色质调控因子的自然平衡被改变时，处于超活性状态的它可能更容易被影响（Lucchesi et al. 2005)。

## 6. 展　望

随着生物芯片（microarray）的广泛利用，我们开始在整个基因组范围内测定 MSL 复合物具体的结合形式和鉴定它所调控的基因（Alekseyenko et al. 2006；Gilfillan et al. 2006；Legube et al. 2006)。这些信息必将使我们更好地了解 MSL 复合物是如何靶向到 X 染色体上的。对靶向性的了解将反过来有助于我们了解 MSL 复合物增加基因表达的机制。最近的研究揭示，是基因的编码区，而不是上游调控区结合 MSL 复合物，正如只有特定的基因的 H4K16 上乙酰化。这个结果为剂量补偿效应可通过调整转录延长率和 RNA 聚合酶的再循环而改变的模型提供了关键的证据（Smith et al. 2001)。

要想充分阐明剂量补偿效应发生的分子机制，还需要精确地确定 MSL 复合物对普

通转录作用的影响。因为 MSL 蛋白和 roX RNA 都不是雌性生殖所需要的，因此它们似乎都致力于剂量补偿效应。也有一些辅助因子利于剂量补偿效应的发生，但它们在通常的核、染色体或转录上的功能是雄性和雌性所必需的（Mendjan et al. 2006）。在将来的几年中，识别这些因子是非常重要的工作，它将增进对 X 染色体连锁基因在雄性果蝇中两倍表达的生化基础。最终，剂量补偿作用将有助于我们了解高等动物中染色质转录转换的关键信息，也将为精确阐明遗传基因和谐表达的分子机制提供重要线索。

吕雯雯 译
王大亮 校
孙方霖 审校

## 参考文献

Aasland R. and Stewart A.F. 1995. The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. *Nucleic Acids Res.* **23:** 3168–3173.

Akhtar A. and Becker P.B. 2000. Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in *Drosophila*. *Mol. Cell* **5:** 367–375.

Alekseyenko A.A., Larschan E., Lai W.R., Park P.J., and Kuroda M.I. 2006. High-resolution ChIP-chip analysis reveals that the *Drosophila* MSL complex selectively identifies active genes on the male X chromosome. *Genes Dev.* **20:** 848–857.

Amrein H. and Axel R. 1997. Genes expressed in neurons of adult male *Drosophila*. *Cell* **88:** 459–469.

Badenhorst P., Voas M., Rebay I., and Wu C. 2002. Biological functions of the ISWI chromatin remodeling complex NURF. *Genes Dev.* **16:** 3186–3198.

Bashaw G.J. and Baker B.S. 1997. The regulation of the *Drosophila* msl-2 gene reveals a function for sex-lethal in translational control. *Cell* **89:** 789–798.

Birchler J.A., Pal-Bhadra M., and Bhadra U. 2003. Dosage dependent gene regulation and the compensation of the X chromosome in *Drosophila* males. *Genetica* **117:** 179–190.

Bone J.R., Lavender J., Richman R., Palmer M.J., Turner B.M., and Kuroda M.I. 1994. Acetylated histone H4 on the male X chromosome is associated with dosage compensation in *Drosophila*. *Genes Dev.* **8:** 96–104.

Cline T.W. and Meyer B.J. 1996. Vive la difference: Males vs females in flies vs worms. *Annu. Rev. Genet.* **30:** 637–702.

Copps K., Richman R., Lyman L.M., Chang K.A., Rampersad-Ammons J., and Kuroda M.I. 1998. Complex formation by the *Drosophila* MSL proteins: Role of the MSL2 RING finger in protein complex assembly. *EMBO J.* **17:** 5409–5417.

Corona D.F., Clapier C.R., Becker P.B., and Tamkun J.W. 2002. Modulation of ISWI function by site-specific histone acetylation. *EMBO Rep.* **3:** 242–247.

Delattre M., Spierer A., Jaquet Y., and Spierer P. 2004. Increased expression of *Drosophila* Su(var)3-7 triggers Su(var)3-9-dependent heterochromatin formation. *J. Cell Sci.* **117:** 6239–6247.

Demakova O.V., Kotlikova I.V., Gordadze P.R., Alekseyenko A.A., Kuroda M.I., and Zhimulev I.F. 2003. The MSL complex levels are critical for its correct targeting to the chromosomes in *Drosophila melanogaster*. *Chromosoma* **112:** 103–115.

Deuring R., Fanti L., Armstrong J.A., Sarte M., Papoulas O., Prestel M., Daubresse G., Verardo M., Moseley S.L., Berloco M., et al. 2000. The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. *Mol. Cell* **5:** 355–365.

Eisen A., Utley R.T., Nourani A., Allard S., Schmidt P., Lane W.S., Lucchesi J.C., and Cote J. 2001. The yeast NuA4 and *Drosophila* MSL complexes contain homologous subunits important for transcription regulation. *J. Biol. Chem.* **276:** 3484–3491.

Fagegaltier D. and Baker B.S. 2004. X chromosome sites autonomously recruit the dosage compensation complex in *Drosophila* males. *PLoS Biol.* **2:** e341.

Gilfillan G.D., Dahlsveen I.K., and Becker P.B. 2004. Lifting a chromosome: Dosage compensation in *Drosophila melanogaster*. *FEBS Lett.* **567:** 8–14.

Gilfillan G.D., Straub T., de Wit E., Greil F., Lamm R., van Steensel B., and Becker P.B. 2006. Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes Dev.* **20:** 858–870.

Gorman M., Franke A., and Baker B.S. 1995. Molecular characterization of the male-specific lethal-3 gene and investigations of the regulation of dosage compensation in *Drosophila*. *Development* **121:** 463–475.

Gu W., Szauter P., and Lucchesi J.C. 1998. Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*. *Dev. Genet.* **22:** 56–64.

Hamada F.N., Park P.J., Gordadze P.R., and Kuroda M.I. 2005. Global regulation of X chromosomal genes by the MSL complex in *Drosophila melanogaster*. *Genes Dev.* **19:** 2289–2294.

Hilfiker A., Hilfiker-Kleiner D., Pannuti A., and Lucchesi J.C. 1997. mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.* **16:** 2054–2060.

Jin Y., Wang Y., Walker D.L., Dong H., Conley C., Johansen J., and Johansen K.M. 1999. JIL-1: A novel chromosomal tandem kinase implicated in transcriptional regulation in *Drosophila*. *Mol. Cell* **4:** 129–135.

Kageyama Y., Mengus G., Gilfillan G., Kennedy H.G., Stuckenholz C., Kelley R.L., and Kuroda M.I. 2001. Association and spreading of the *Drosophila* dosage compensation complex from a discrete roX1 chromatin entry site. *EMBO J.* **20:** 2236–2245.

Kelley R.L. 2004. Path to equality strewn with roX. *Dev. Biol.* **269:** 18–25.

Kelley R.L., Wang J., Bell L., and Kuroda M.I. 1997. Sex lethal controls dosage compensation in *Drosophila* by a non-splicing mechanism. *Nature* **387:** 195–199.

Kelley R.L., Meller V.H., Gordadze P.R., Roman G., Davis R.L., and Kuroda M.I. 1999. Epigenetic spreading of the *Drosophila* dosage compensation complex from roX RNA genes into flanking chromatin. *Cell* **98:** 513–522.

Kelley R.L., Solovyeva I., Lyman L.M., Richman R., Solovyev V., and Kuroda M.I. 1995. Expression of msl-2 causes assembly of dosage compensation regulators on the X chromosomes and female lethality in *Drosophila*. *Cell* **81:** 867–877.

Koonin E.V., Zhou S., and Lucchesi J.C. 1995. The chromo superfamily New members, duplication of the chromo domain and possibl role in delivering transcription regulators to chromatin. *Nuclei Acids Res.* **23:** 4229–4233.

Kuroda M.I., Kernan M.J., Kreber R., Ganetzky B., and Baker B.S. 1991 The maleless protein associates with the X chromosome to regulat dosage compensation in *Drosophila*. *Cell* **66:** 935–947.

Kusch T., Florens L., Macdonald W.H., Swanson S.K., Glaser R.L., Yate J.R., III, Abmayr S.M., Washburn M.P., and Workman J.L. 2004 Acetylation by Tip60 is required for selective histone varian exchange at DNA lesions. *Science* **306:** 2084–2087.

Lee C.C., Chang K.A., Kurtoda M.I., and Hurwitz J. 1997. Th NTPase/helicase activities of *Drosophila* maleless, an essential fac tor in dosage compensation. *EMBO J.* **16:** 2671–2681.

Legube G., McWeeney S.K., Lercher M.J., and Akhtar A. 2006. X-chro mosome-wide profiling of MSL-1 distribution and dosage com pensation in *Drosophila*. *Genes Dev.* **20:** 871–883.

Lucchesi J.C. and Manning J.E. 1987. Gene dosage compensation i *Drosophila melanogaster*. *Adv. Genet.* **24:** 371–429.

Lucchesi J.C., Kelly W.G., and Panning B. 2005. Chromatin remodeling in dosage compensation. *Annu. Rev. Genet.* **39:** 615–651.

Marin I. 2003. Evolution of chromatin-remodeling complexes: Comparative genomics reveals the ancient origin of "novel" compensasome genes. *J. Mol. Evol.* **56:** 527–539.

Meller V.H. and Rattner B.P. 2002. The roX genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. *EMBO J.* **21:** 1084–1091.

Meller V.H., Wu K.H., Roman G., Kuroda M.I., and Davis R.L. 1997. roX1 RNA paints the X chromosome of male *Drosophila* and is regulated by the dosage compensation system. *Cell* **88:** 445–457.

Meller V.H., Gordadze P.R., Park Y., Chu X., Stuckenholz C., Kelley R.L., and Kuroda M.I. 2000. Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in *Drosophila*. *Curr. Biol.* **10:** 136–143.

Mendjan S., Taipale M., Kind J., Holz H., Gebhardt P., Schelder M., Vermeulen M., Buscaino A., Duncan K., Mueller J., et al. 2006. Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol. Cell* **21:** 811–823.

Mito Y., Henikoff J.G., and Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* **37:** 1090–1097.

Morales V., Straub T., Neumann M.F., Mengus G., Akhtar A., and Becker P.B. 2004. Functional integration of the histone acetyltransferase MOF into the dosage compensation complex. *EMBO J.* **23:** 2258–2268.

Morgan T.H. 1932. *The scientific basis of evolution*. Norton, New York, p. 80.

Oh H., Bone J.R., and Kuroda M.I. 2004. Multiple classes of MSL binding sites target dosage compensation to the X chromosome of *Drosophila*. *Curr. Biol.* **14:** 481–487.

Oh H., Park Y., and Kuroda M.I. 2003. Local spreading of MSL complexes from roX genes on the *Drosophila* X chromosome. *Genes Dev.* **17:** 1334–1339.

Palmer M.J., Mergner V.A., Richman R., Manning J.E., Kuroda M.I., and Lucchesi J.C. 1993. The male-specific lethal-one (msl-1) gene of *Drosophila melanogaster* encodes a novel protein that associates with the X chromosome in males. *Genetics* **134:** 545–557.

Pannuti A. and Lucchesi J.C. 2000. Recycling to remodel: Evolution of dosage-compensation complexes. *Curr. Opin. Genet. Dev.* **10:** 644–650.

Pardo P.S., Leung J.K., Lucchesi J.C., and Pereira-Smith O.M. 2002. MRG15, a novel chromodomain protein, is present in two distinct multiprotein complexes involved in transcriptional activation. *J. Biol. Chem.* **277:** 50860–50866.

Park Y., Kelley R.L., Oh H., Kuroda M.I., and Meller V.H. 2002. Extent of chromatin spreading determined by roX RNA recruitment of MSL proteins. *Science* **298:** 1620–1623.

Sass G.L., Pannuti A., and Lucchesi J.C. 2003. Male-specific lethal complex of *Drosophila* targets activated regions of the X chromosome for chromatin remodeling. *Proc. Natl. Acad. Sci.* **100:** 8287–8291.

Schalch T., Duda S., Sargent D.F., and Richmond T.J. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436:** 138–141.

Scott M.J., Pan L.L., Cleland S.B., Knox A.L., and Heinrich J. 2000. MSL1 plays a central role in assembly of the MSL complex, essential for dosage compensation in *Drosophila*. *EMBO J.* **19:** 144–155.

Sif S. 2004. ATP-dependent nucleosome remodeling complexes: Enzymes tailored to deal with chromatin. *J. Cell Biochem.* **91:** 1087–1098.

Smith C.L. and Peterson C.L. 2005. ATP-dependent chromatin remodeling. *Curr. Top. Dev. Biol.* **65:** 115–148.

Smith E.R., Allis C.D., and Lucchesi J.C. 2001. Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in *Drosophila* males. *J. Biol.Chem.* **276:** 31483–31486.

Smith E.R., Cayrou C., Huang R., Lane W.S., Cote J., and Lucchesi J.C. 2005. A human protein complex homologous to the *Drosophila* MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. *Mol. Cell. Biol.* **25:** 9175–9188.

Smith E.R., Pannuti A., Gu W., Steurnagel A., Cook R.G., Allis C.D., and Lucchesi J.C. 2000. The *Drosophila* MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. *Mol. Cell. Biol.* **20:** 312–318.

Smith E.R., Eisen A., Gu W., Sattah M., Pannuti A., Zhou J., Cook R.G., Lucchesi J.C., and Allis C.D. 1998. ESA1 is a histone acetyltransferase that is essential for growth in yeast. *Proc. Natl. Acad. Sci.* **95:** 3561–3565.

Straub T., Dahlsveen I.K., and Becker P.B. 2005a. Dosage compensation in flies: Mechanism, models, mystery. *FEBS Lett.* **579:** 3258–3263.

Straub T., Gilfillan G.D., Maier V.K., and Becker P.B. 2005b. The *Drosophila* MSL complex activates the transcription of target genes. *Genes Dev.* **19:** 2284–2288.

Straub T., Neumann M.F., Prestel M., Kremmer E., Kaether C., Haass C., and Becker P.B. 2005c. Stable chromosomal association of MSL2 defines a dosage-compensated nuclear compartment. *Chromosoma* **114:** 352–364.

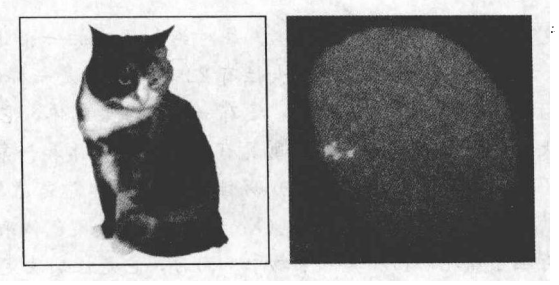
Stuckenholz C., Meller V.H., and Kuroda M.I. 2003. Functional redundancy within roX1, a noncoding RNA involved in dosage compensation in *Drosophila melanogaster*. *Genetics* **164:** 1003–1014.

Suka N., Luo K., and Grunstein M. 2002. Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat. Genet.* **32:** 378–383.

Turner B.M., Birley A.J., and Lavender J. 1992. Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. *Cell* **69:** 375–384.

Wang Y., Zhang W., Jin Y., Johansen J., and Johansen K.M. 2001. The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in *Drosophila*. *Cell* **105:** 433–443.

# 第 17 章



# 哺乳动物中的剂量补偿效应

Neil Brockdorff[1] and Bryan M. Turner[2]

[1]*X Inactivation Group (Developmental Epigenetics Group), MRC Clinical Sciences Centre, ISCM Hammersmith Campus, London W12 ONN, United Kingdom*

[2]*Chromatin and Gene Expression Group, Institute of Biomedical Research, University of Birmingham Medical School, Birmingham B15 2TT, United Kingdom*

## 本章目录

1. 序言
   - 1.1 有性生殖的优势
   - 1.2 性别决定的方法
   - 1.3 性别决定的染色体方法提出剂量补偿的需求
   - 1.4 雌性哺乳动物中失活的 X 染色体的识别
2. 重要概念
   - 2.1 X 染色体失活受发育调控
   - 2.2 染色体沉默包括多层次的表观遗传修饰
   - 2.3 一些基因逃避 X 染色体失活
   - 2.4 X 染色体的失活受一个重要的开关部位：X 染色体的失活中心的调控
3. X 染色体失活的起始
   - 3.1 X 染色体的印记失活和随机失活
   - 3.2 X 染色体印记失活的调控
   - 3.3 X 染色体的随机失活的调控——计数
   - 3.4 X 染色体的随机失活的调控——选择
   - 3.5 早期胚胎基因失活的转换模式
4. 失活状态的增殖和维持
   - 4.1 *Xist* RNA、基因沉默和异染色质组装
   - 4.2 失活 X 染色体的异染色质结构
   - 4.3 X 染色体失活的组蛋白修饰酶学
   - 4.4 ES 细胞作为模式系统研究导致 X 染色体失活的有序过程
   - 4.5 沉默染色质的蔓延
   - 4.6 X 染色质逃避失活
   - 4.7 袋类哺乳动物中的 X 染色体失活
5. X 染色体的再激活和重编程
   - 5.1 体细胞中 X 染色体失活的稳定性
   - 5.2 正常发育中 X 染色体的再激活
   - 5.3 重编程实验中 X 染色体的再激活
   - 5.4 诱导 *Xist* 转基因表达的启示
6. 总结和展望

参考文献

## 概要

人类的DNA组装成23对大小不同的染色体。每对染色体中的一条遗传自父亲（父本同系物），另一条则来自母亲（母本同系物）。男性和女性有22对染色体是一样的，称为常染色体，它们按照个体由大到小的顺序编为1～22号。此外，还有一对为性染色体，它们因性别差异而不同。女性有两条一样的中等大小的X染色体；而男性则具有一条X染色体和一条Y染色体。较之X染色体，Y染色体个体偏小，且携带的基因也较少。男性的X染色体总是遗传自母亲，Y染色体遗传自父亲。同样，女性的一条X染色体来自母亲（Xm），另一条X染色体来自于父亲（Xp）。这种性别间染色体差异普遍存在于哺乳动物和其他生物中。同时，它是性别决定的生物机制的一部分。但是，它也引发了一些进化上的问题：两性所携带的X连锁基因的数量不同，女性是男性的两倍。这会导致基因产物（RNA和蛋白）数量上的不平衡，从而要求代谢调控和其他细胞进程的差异。为避免这一现象的发生，剂量补偿机制进化到用来平衡两性间X连锁基因产物。

在哺乳动物中，剂量补偿的机制包括关闭（沉默）两条X染色体中一条的大部分基因，使雌性同雄性一样，仅有一条激活的X染色体。这种调控方式，通常称为X染色体失活，是1961年由Mary Lyon首先提出，用来解释小鼠中X染色体连锁的毛色基因（coat-color gene）的表达形式与斑猫（calico cat）显示的毛色形式是类似的。自此之后40多年的大量研究一直致力于分析这个过程中有趣而复杂的机制。已知X染色体失活发生在发育的早期，但机制却很复杂。在雌性动物发育早期，当胚胎仅有少数细胞组成时，父本的X染色体在所有细胞中选择性失活。Xp必须以某种方式被标注“印记”而失活。后来，在胚泡期（blastocyst）（刚好在植入前），当胚胎由50～100个细胞组成时，在这些细胞将继续形成胚胎（定位在内细胞团内，inner cell mass，ICM）的细胞中，Xp再激活。因此，雌性中有两条激活的X染色体。再后来，Xp或Xm被选择性随机失活，所携带的基因被沉默。有趣的是，胚泡中将形成胚胎外组织（胎盘和卵黄囊）的细胞内，Xp仍然沉默。内细胞团中，选择哪条X染色体沉默仍是一个让人迷惑不解的问题。

失活的X染色体在后代细胞中仍然保持沉默。这是我们所了解的基因沉默最稳定的形式之一。我们试图用实验来逆转这一现象，却一直没有成功。然而，在雌性的生殖细胞——卵母细胞中，可以通过减数分裂来逆转失活的过程，使之拥有两条激活的X染色体，在其成熟后又变为一条，单倍体卵（haploid ovum）也被激活。

X染色体失活过程的研究揭示了基因沉默的新的分子机制。雌性X染色体上有被称为*XIST*的基因，由其中一条*X*染色体上*XIST*基因转录而来的非编码RNA表达的增高驱动了沉默的起始。正如细胞核绿染区域所显示，这些RNA包裹着含有处于开放状态的*XIST*基因的X染色体。之后，起始了染色体上所有基因的沉默。*XIST*基因本身维持开放状态。*XIST*包被后，失活并沉默的X染色体经过一系列的改变。主要的DNA组装蛋白——组蛋白在其功能重要位点发生了一系列化学修饰。例如，某些赖氨酸残基乙酰化水平迅速降低，而某些赖氨酸的甲基化水平升高。这些变化是由靶向沉默的X染色体的特异性酶来执行的。另外，在失活的X染色体上，组蛋白的变体macro-H2A代替一部分正常的H2A。随后，在失活的X染色体上的特定区域发生DNA甲基化。染色体失活（X inactivation，Xi）是一个与基因长期沉默相关的过程。这些变化使X染色体具有一个致密的特别结构，在细胞核内是一个可见的致密DNA的特征性的斑点，现在被称为巴氏小体（Barr body）。

近年来，X染色体失活的研究已为基因沉默的主要表观遗传学机制和发育过程中基因表达的调控提供了重要的线索。我们可以肯定地预见到这一研究的发展。

# 1. 序　言

## 1.1　有性生殖的优势

有性生殖在真核生物中普遍存在，即便以扦插或营养生殖方式完成繁殖的植物也常具有可供选择的有性生殖模式。很多时候，这一现象被认为是进化的结果。有性生殖引发了大量的遗传变异，在此基础上发生自然选择，从而进化。每次有性生殖都伴随着等位基因的重组，这种方式产生的个体较之无性繁殖产生的个体，能够更有效地应对环境的改变。但是，高等真核生物的性别差异是复杂的，它需要能够产生雄性和雌性的生殖器官的发育途径、进行减数分裂需要的生理学和生物学结构、生殖细胞的成熟、配偶体的吸引和交配（这些问题的进一步讨论，参考 Marshall Graves and Shetty 2001 及其中的参考文献）。非常关键的一点，也是衡量进化成功与否的明确方法，是多样化群体是否能够更好地避免最终的灭绝之灾。

## 1.2　性别决定的方法

生物体间决定不同性别的遗传机制差异很大。最简单的系统由单基因座，即在一种性别中是纯合子（同配性别），而另一种性别中是杂合子（异配性别）（图 17-1）。这一简单系统以不同的方式进化，达到不同生物体中不同复杂程度的要求。在一些生物体中，抑制异配性别中决定性别等位基因的减数分裂重组（染色体交叉，cross-over）的机制已经成立（图 17-1），这一步可以有助于防止子代出现等位基因混淆而导致的雌雄间性现象。在许多情况下，一条染色体的部分或全部无法重组，并伴随着遗传信息的丢失。促使这条染色体退化的进化原因仍不清楚。但在许多物种中最终的结果是两性的差别不仅仅在等位基因的一个或几个位点，而是整条染色体的差别。正如表 17-1 中简单的示意图所示。在大多数物种中，包括我们人类自己，是雄性携带退化的染色体，即使存在一些例外。

在发育过程中，性别分化通常是由一个或少数关键基因表达或不表达引起的。这些基因的产物引起一系列基因调控过程的发生，下调了一条或其他性别决定的通路（线虫的详情参考第 15 章和果蝇的详情参考第 16 章）。在人类，Y 染色体上 *SRY* 基因的蛋白产物能够使胚胎进入男性的性别发育通路（参考综述部分 McElreavey 1995）。这种机制的顺利进行并不需要大量染色体的差别，那么为什么这种差别如此常见？这可能是交换染色体抑制过程的副产物，可防止雌雄间性的产生（图 17-1）。对影响种群中等位基因扩展的因素的数学分析表明，抑制交换必然导致有害突变的积累（突变甚至可能引起交换染色体的进一步抑制）。这将导致最初两条同源染色体的一条逐渐退化（图 17-1）。这个过程被定义为“Muller 棘轮”（Muller's ratchet），以纪念第一个提出它的遗传学家。只能这样认为，这只是最初采取双性策略繁殖的结果（进一步的讨论请参考 Charlesworth 1996 及参考文献）。不管染色体退化背后是怎样的进化驱动，染色体退化的事实（推测可能继续发生）需要协同的进化机制来应对同种物种成员间主要染色体差异。这种机制被称为是剂量补偿机制。

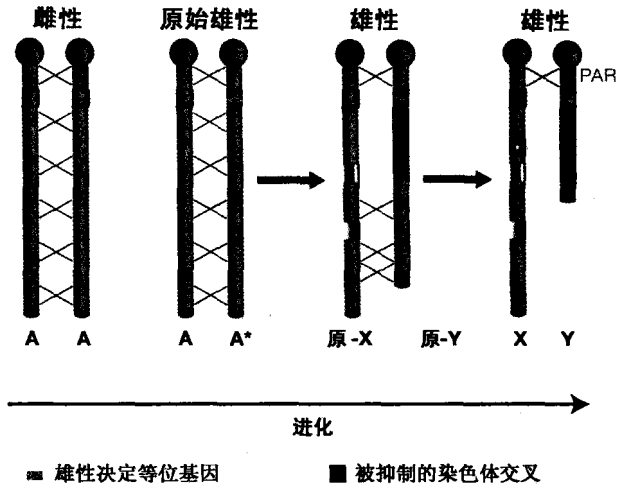



图 17-1 Y染色体的进化

在进化的早期，两性间的差别可能只在单个常染色体基因座上：一种性别（原始的雄性，proto-male）在这个基因座上是杂合的，另一种性别（称为原始的雌性）是纯合的。“雄性决定等位基因”用黄色表示。如果交配需要每个性别的一条染色体，那么个体就不会产生个体雄性等位基因的纯合子。在这个早期，与酵母中两种交配的差异相比，两性间生理的差异将非常精细。为了防止雌雄兼性的情况形成，染色体交换将在雄性决定基因座中，或其周围被抑制（抑制的区域用深颜色标出）。由于交换抑制将减少种群扩展的能力，突变将在这个区域内积累，因此，选择的压力抑制了它们。染色体交换被抑制的退化区域逐渐增大（“Muller的棘轮”学说），直到这条染色体丢失它大部分的活性和功能基因。为了在有丝分裂期间染色体配对和交换，Y染色体上必须保留一小段活性区域和X染色体同源。这就是伪常染色体区（pseudoautosomal region，PAR），最初的常染色体进化之后，相对应的Y染色体大部分是通过其他染色体上的异位，最终形成特征性的X染色体。和其他染色体一样，X染色体是通过进化把不同阶段的DNA片段镶嵌到合适的位置，它们中一些是古代的，一些是相对近代的。这在原始X染色体和X染色体上通过不同的拼凑形式展示。人类的X染色体则富集了逃脱了X染色体失活的相对近代的基因。

## 1.3 性别决定的染色体方法提出剂量补偿的需求

在哺乳动物和果蝇中，雄性有每条性染色体的一个拷贝：一条X染色体和一条Y染色体。而雌性有两条X染色体。这两个种群中，Y染色体上的基因很少，且大部分是异染色质。Y染色体仅有一些雄性发育和生育所必需的基因。相反，X染色体个体较大，并携带大量的基因。如果对错误置之不理，拷贝数目上的双倍差异将导致性别间几千种基因产物细胞内浓度的双倍差异。显然，进化不允许相同物种成员间有如此大的差异。目前，已经发育出通过剂量补偿效应机制消除这种差异的方法。

通常，仅有三种方式使X连锁基因转录水平在同型配子和异型配子中相同。它们分别是：①关闭两条雌性X染色体中一条上的基因；②使雄性单条X染色体的基因转录率加倍；③使每条雌性X染色体的转录率减半。后两种方式的最后结果一样：雄性X染色体上基因的转录速度是雌性两条X染色体同样基因转录速度的两倍，但达到这种情况的路径却完全不同。三种机制的例子已经证实。哺乳动物是第一种机制，果蝇是

第二种（见第 16 章），线虫为第三种（第 15 章）。剂量补偿效应中，这三种显著不同的机制已经进化到相对独立，这证明性别间 X 连锁基因产物水平相同的最终的目标非常重要（Marin et al. 2000）。因此，破坏这三种方式中的任何一种剂量补偿效应的突变都可使受影响的性别致死。

### 1.4　雌性哺乳动物中失活的 X 染色体的识别

1949 年，Barr 和 Bertram 描述的性染色质体是存在于不同哺乳动物物种的雌性（而不是雄性）细胞核中（参考题图）的一种光镜下可见的结构。这个结构有益于性发育异常的研究。但是直到 1959 年，Ohno 和他的同事才证明这个结构来自于雌性两条 X 染色体中的一条。1961 年，Mary Lyon 描述了雌性小鼠中 X 染色体连锁的毛色基因表达的遗传实验。为了解释个别雌性小鼠中可变混合毛色（嵌合体）的遗传形式，Lyon 假设在每个雌性细胞中，两条雌性 X 染色体的一条在发育的早期稳定地失活（Lyon 1961）。性染色质体，现在称为巴氏小体（barr body），是失活 X 染色体的细胞学表现。使用雌性杂合子皮肤成纤维细胞进行 X 染色体连锁酶 6-磷酸葡萄糖脱氢酶（glucose-6-phosphate dehydrogenase，G6PD）多态性的实验，结果显示两个等位基因中仅有一个在单细胞克隆（克隆）中表达。因此，表明了从一代细胞到下一代的失活状态的可遗传性（Davidson et al. 1963），也证实了女性中 X 染色体失活的发生（Beutler et al. 1962）。通过用多重拷贝 X 染色体（例如 47XXX 或 48XXXX 染色体组型）对女性中 X 染色体失活的进一步研究表明，所有的 X 染色体中，除一条以外，其余所有的均失活。这被总结为“$n-1$ 规则”，即，如果一个个体有 $n$ 条 $X$ 染色体，那么就有 $n-1$ 条将要被失活（Ohno，1967）。这个规则解释了 X 染色体非整倍体非常轻微的临床症状。X 染色体失活的假说继续解释雌性细胞中 X 染色体连锁基因表达的特性，从首次提出以来一直没有本质的改变。过去的约 40 年里，人们一直试图找到这一现象发生的分子机制。

## 2. 重要概念

### 2.1　X 染色体失活受发育调控

在哺乳动物中，雌性 X 染色体失活受发育的调控。早期受精卵（zygote）中两条 X 染色体都是活化的（Epstein et al. 1978），它们是在细胞分化过程中失活的。通常，细胞中来源于母本或父本的 X 染色体（分别 Xm 或 Xp）失活的几率是相等的。有一种情况例外，只有印记的 X 染色体失活，它发生于袋类动物的一生和小鼠早期着床前胚胎，通常是 Xp 失活。在后一种情况中，印记的 Xp 染色体失活维持在分化的第一谱系，称为胚外滋养外胚层（extraembryonic trophectoderm，TE）细胞和原始内胚层（primitive endoderm cells，PE）细胞，但在产生胚胎内细胞团细胞时，失活的 X 染色体可以被再激活。失活 X 染色体再激活的现象也发生在发育中的原生殖细胞（primordial germ cell，PGC）中，这就确保在配子中 X 染色体再激活。图 17-2 展示雌鼠中 X 染色体失活和再激活的周期。

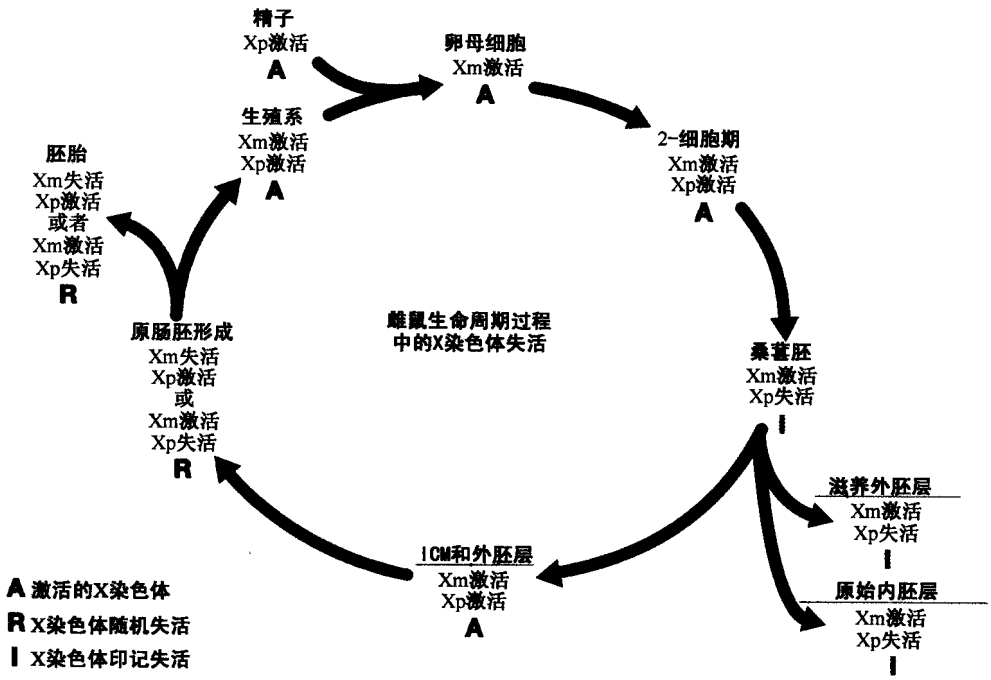



图 17-2 X 染色体失活和再激活的循环

在发育过程中 X 染色体经历失活和再激活的循环。红色箭头表示 X 染色体失活的阶段，绿色的箭头表示 X 染色体再激活的阶段。失活首先发生在早期植入前胚胎（X 染色体的印记失活），然后在原肠胚形成时发生在外胚层的细胞内（X 染色体的随机失活）。失活的 X 染色体在内细胞团（ICM）内囊泡期它们第一次被定位时或是在发育的生殖细胞里被重新激活。

## 2.2 染色体沉默包括多层次的表观遗传修饰

X 染色体的沉默是在染色体结构水平上通过组蛋白尾部修饰完成的，结合或去除不同的组蛋白和 CpG 岛 DNA 甲基化均有助于形成稳定的异染色质结构。在个体发育的过程中，各层次表观遗传修饰是逐步建立的，这在 4.4 节将详述。整体而言，多层次的表观遗传修饰确保失活的 X 染色体在多次细胞分裂过程中稳定繁殖。

个体修饰的作用在不同谱系中的变化是短期的。例如，在早期发育中，X 失活需要高水平的 H3K27me3（组蛋白 H3 在 27 位赖氨酸的三甲基化），而后期发育则不需要。相反，CpG 岛甲基化仅在后期发育中是必须的，而且在滋养外胚层细胞的发育中完全不需要。

## 2.3 一些基因逃避 X 染色体失活

X 染色体失活会影响 X 染色体上的大部分基因，但也有一些基因可逃避沉默。其中包括与雄性减数分裂中与 Y 染色体配对的 X 染色体上一小部分区域，这部分区域称为拟常染色体区域（pseudoautosomal region，PAR）或 XY 配对区域（图 17-1）。因为雄性和雌性中均有这些基因的双拷贝，所以这部分区域里的基因不需要剂量补偿效应。

其他逃避沉默的基因，还包括具有和没有 Y 染色体连锁同系物的基因。最近的研究表明，人类 X 染色体上大约有 15%的基因逃避 X 染色体失活（Carrel and Willard 2005）。有趣的是，一些基因位于 Xp 染色体的短臂上，在进化时间上，这些基因是相对最近才有的。小鼠的研究表明，在个体发育早期，这些基因起初失活，它们是在发育过程中逐渐被激活的（4.6 节）。袋类动物的研究表明，一些基因在一定程度上逃避 X 染色体失活。它反映出整个个体发育过程中沉默维持的失败，这可能与一些物种 Xi 缺少 CpG 岛甲基化有关（详见 4.7 节）。

## 2.4　X 染色体的失活受一个重要的开关部位：X 染色体的失活中心的调控

传统的遗传学研究表明，X 染色体失活受单一的顺式作用总开关位点的控制，即 X 染色体失活中心（X inactive center，Xic）。X 染色体顺式沉默需要这个失活中心，它确保 X 染色体随机失活起始的正确性和适当性。最近，有更多在分子水平上对 Xic 的特征的研究。这个基因座产生一个大的非编码 RNA 称为 *Xist*（X 染色体失活特异性转录，*X inactive specific transcript*），它具有顺式结合的特点，并从转录位点开始沿着整条染色体积累（图 17-3）（Brown et al. 1991，1992；BrockDorff et al. 1992）。覆盖有 *Xist*

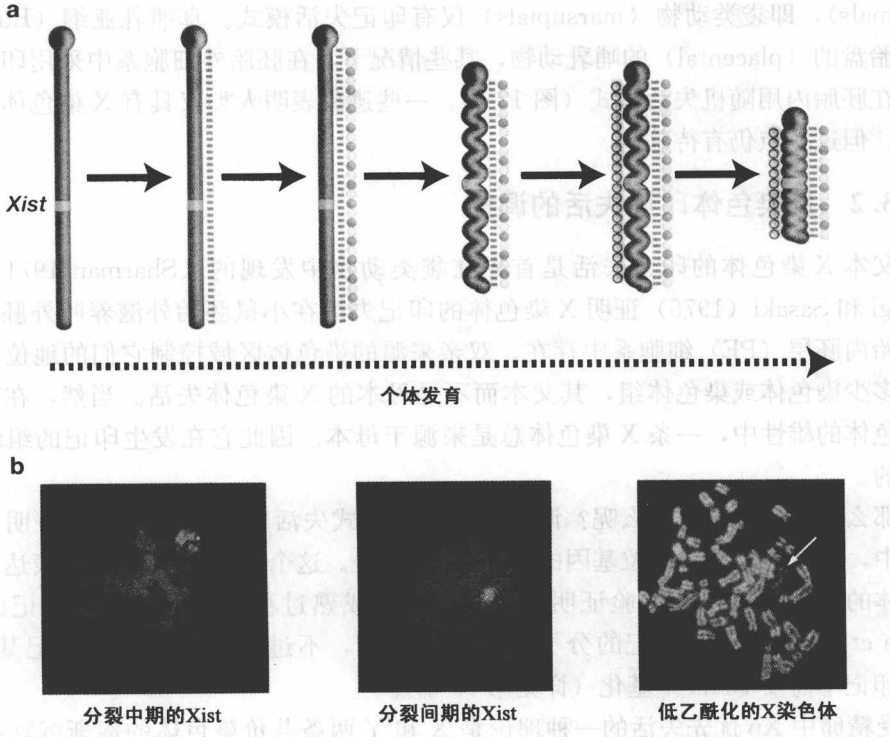



图 17-3　*Xist* RNA 诱导染色体范围逐渐异染色质化

当 *Xist* RNA 表达时，其 RNA 从它的转录位点（绿色虚线）开始结合并包围 X 染色体上。*Xist* RNA 通过招募染色体修饰活性开启染色体沉默（红色和黄色的圆球）。最初沉默的波动可招募其他层次表观遗传修饰（灰色的圆球），进一步稳定异染色质的结构。通过发育和个体发生逐步建立不同水平的表观遗传沉默。间期和中期中，沿着 X 染色体上 *Xist* RNA 的原位杂交定位。

RNA 的染色体为 X 染色体沉默的启动提供了准备（Lee et al. 1996；Penny et al. 1996；Wutz and Jaenisch 2000）。到目前的研究表明，这一过程至少部分是通过 *Xist* 介导的染色体修饰复合物的募集发生的。

另一个非编码 RNA，*Tsix*，也定位在 Xic 区域（Lee et al. 1999），在 *Xist* 表达调控中起到关键作用。*Tsix* 与 *Xist* 基因重叠，但却反向转录。因此，它的名字 *Tsix*，是 *Xist* 的反向拼写。

## 3. X 染色体失活的起始

### 3.1 X 染色体的印记失活和随机失活

决定 X 染色体失活需要密切的调控。雄性细胞必须避免其唯一的 X 染色体沉默，雌性细胞也需要避免两条 X 染色体都失活或都活化。已经证明，有两种不同的调控模式。X 染色体失活印记模式沉默来源于父本 X 染色体。而在染色体失活随机模式中，来源于父本和母本的 X 染色体失活几率是相同的。后兽亚纲的哺乳动物（metatherian mammals），即袋类动物（marsupials）仅有印记失活模式。真哺乳亚纲（Eutherian），即有胎盘的（placental）的哺乳动物，某些情况下，在胚胎外细胞系中采用印记失活模式；在胚胎内用随机失活模式（图 17-2）。一些迹象表明人类仅具有 X 染色体失活随机模式，但这一点仍有待确定。

### 3.2 X 染色体印记失活的调控

父本 X 染色体的印记失活是首先在袋类动物中发现的（Sharman 1971）。后来，Takagi 和 Sasaki（1975）证明 X 染色体的印记失活在小鼠胚胎外滋养叶外胚层（TE）和原始内胚层（PE）细胞系中存在。双亲来源的染色体区域控制它们的地位，即不管出现多少染色体或染色体组，其父本而不是母本的 X 染色体失活。当然，在有 XY 两条染色体的雄性中，一条 X 染色体总是来源于母本。因此它在发生印记的组织中是不失活的。

那么印记的本质是什么呢？调节 X 染色体顺式失活的 *Xist* 基因研究表明，在小鼠胚胎中，存在抑制 Xm 等位基因的印记（图 17-4）。这个印记阻止 *Xist* 的表达，保持 X 染色体的活性。核移植实验证明，在卵母细胞成熟过程中，*Xist* 抑制印记已经建立（Tada et al. 2000）。该印记的分子基础还不清楚，不过与许多其他的印记基因不同，*Xist* 印记不需要 DNA 甲基化（详见第 19 章）。

受精卵中 Xp 优先失活的一种理论是 X 和 Y 两条共价染色体的逐渐沉默在雄性减数分裂的粗线期已经完成（减数分裂的性染色体失活，meiotic sex chromosome inactivation，MSCI）（Huynh and Lee 2003）。最近的研究反对这种观点：第一，已经证明 MSCI 有一个独立于 *Xist* 的机制，它通过粗线期性染色体和常染色体上不配对染色体区域的而开启（Turner et al. 2005 及其参考文献）。第二，早期受精卵中一些 X 染色体连锁基因表达分析显示，在受精卵 Xp 中 *Xist* 的表达下，Xp 的沉默重新发生。

在受精卵基因激活的起始，父本 *Xist* 开始表达。这表明 Xp 中 *Xist* 等位基因是平衡表达的（图 17-4）。然而，在雄性体细胞组织中 *Xist* 的表达被抑制，这就使雄性的生殖细胞必须以一定的方式重塑 *Xist* 的基因座。与此一致的是，在精子发生过程中 *Xist* 的启动子有一个特定区域的 CpG 位点去甲基化（Norris et al. 1994）。胚胎中存在 X 染色体的非整倍体形式，在早期发育过程中，父本的印记控制 *Xist* 的表达，引起 X 染色体不正确的失活形式。这种现象也发生在雄性单性生殖或孤雌生殖的胚胎中，这些胚胎的两条染色体都来源于父本或母本。尽管这不利于胚胎的存活，但有些胚胎仍能存活。显然，X 染色体不当失活的纠正确保了在所有细胞中都有单条激活的 X 染色体。这被认为存在一种超越印记失活的机制，该机制在正常状态下调控 X 染色体的随机失活。

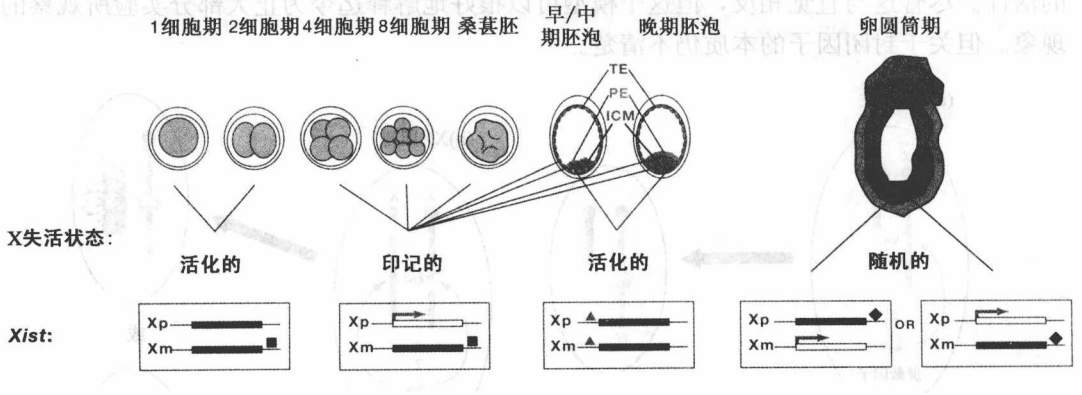



图 17-4　早期发育中 *Xist* 基因的调控

这个图展示在早期 XX 小鼠胚胎中，印记和随机 *Xist* 调控的现有认识和模型。Xm *Xist* 等位基因随受精卵带来一个抑制的标记，可能通过反义的 *Tsix* 基因座介导（黑框）。在 2 细胞阶段胚胎基因活化时，Xp 等位基因准备活化和表达。从 2 细胞形成到桑葚胚期，Xp *Xist* 在所有细胞中表达（白色矩形和 5′端的箭头）。这种表达模式维持在早期胚泡期和后来的 TE、PE 细胞中以及它们完全分化衍生的组织。在胚泡期后期的 ICM 中，*Xist* 的表达消失，可能有 ICM 特异的抑制因子（蓝色的三角形）。然后，在原肠胚形成后，*Xist* 开始表达。屏蔽因子（黑色的菱形）确保 *Xist* 的表达不在两个等位基因中的一条发生（计数）。

*Tsix* 基因是 *Xist* 基因的反义调控子，它是 X 染色体印记失活所必需的。当父本或母本信息传递时，敲除 *Tsix* 的主要启动子可导致胚胎早期致死（Lee 2000）。致死是由于 Xm 中 *Xist* 基因的不适当表达，即 XmY 和 XmXp 在胚胎中保留激活的失败。现在，对于 *Xist* RNA 的表达就是最初的印记或是后来起维持印记的功能仍不清楚。

## 3.3　X 染色体的随机失活的调控——计数

在 X 染色体随机失活的模型中，细胞利用 $n-1$ 的原则，使每个二倍染色体组中除一条 X 染色体外的所有 X 染色体都失活。尽管受到一些因素的影响，但选择哪条染色体沉默（或保持活性）基本上是随机的（参考 3.4 部分）。X 染色体随机失活中，调控子 *Xist* 受到严密的调控。研究表明单个 XX 的胚胎干细胞（embryonic stem，ES）只有表达单个 *Xist* 等位基因时才被分化，但是 XY 细胞从不表达 *Xist*。

在非整倍或多倍 X 染色体的小鼠胚胎中，X 染色体随机失活和印记失活的结果是

不同的。X0 的胚胎是最简单的例子。在印记失活的 X 染色体中，Xm0 的胚胎发育正常，而 Xp0 的胚胎却发育迟缓。这是因为细胞试图失活它的单条 X 染色体，从而减缓胚外组织的发育。在随机 X 染色体失活中，细胞计数 X 染色体的数量（遵循 $n-1$ 的原则），并维持一条 X 染色体的活性，这在 Xm 或 Xp 中均不受影响。

X 染色体失活的随机模式需要细胞有一种检测出现的 X 染色体数量的方法，这通常称为“计数”。Rastan 的封闭因子模型是一个公认的解释计数的模型（Rastan 1983）。这个模型是：在所有细胞中，单个 *Xic*（*Xist* 等位基因）是封闭的，这就指定了活化的 X 染色体。另外 *Xic*（*Xist* 等位基因）的表达诱导 X 染色体的失活（图17-5）。Rastan 的封闭因子模型暗示 X 染色体的失活是一种缺省状态，包括维持一条 X 染色体的活性。尽管这与直觉相反，但这个模型可以很好地解释迄今为止大部分实验所观察的现象。但关于封闭因子的本质仍不清楚。

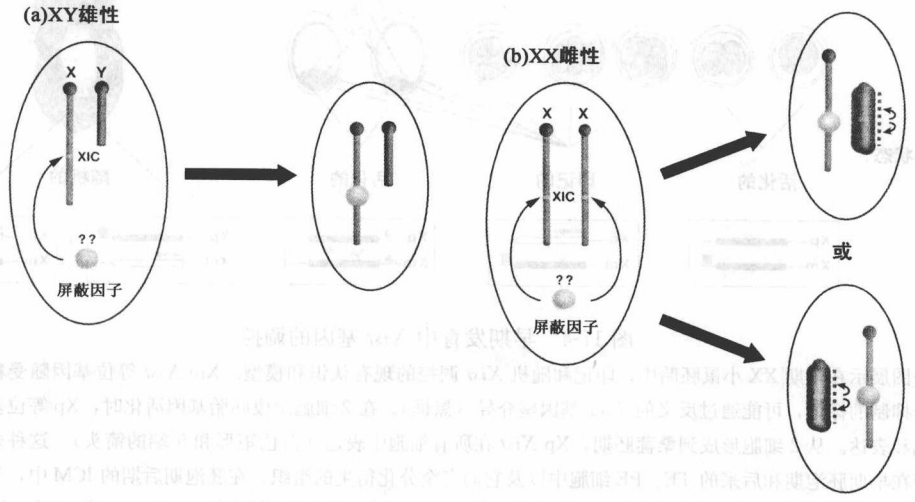



图 17-5 X 染色体失活的屏蔽因子模型

屏蔽因子模型认为有一个因子（黄球）在所有细胞中屏蔽一条染色体上 *Xist* 等位基因（绿框）的表达，以确保有一条染色体不经历 X 染色体失活的过程。(a) 在雄性细胞中，只出现一条等位基因，屏蔽因子总是与之结合。(b) 在雌性细胞中，屏蔽因子结合在 Xm *Xist* 等位基因（红色）和 Xp *Xist* 等位基因（蓝色）的几率是相等的。在 X 染色体失活的开始，只有没有屏蔽的等位基因表达 *Xist* RNA（绿色的虚线）。X 染色体失活中心的不同等位基因可能对屏蔽因子有不同的亲和力，这使杂合的雌性中这个因子更喜欢结合在一条 X 染色体上。在一些情况下，这可能成为最初 X 染色体不随机失活的基础（图 17-7）。

体外胚胎干细胞的培养大大地帮助了我们了解调控 *Xist* 表达和 X 染色体失活的机制。胚胎干细胞来自于胚泡期的小鼠胚胎，尤其是内细胞团（图 17-4）。在添加了白血病抑制因子（leukemia inhibitory factor，LIF）的血清中培养时，细胞可处于未分化状态。从培养基中去除 LIF，并将细胞接种到非贴壁培养板中后，细胞则开始体外分化。在这种条件下，细胞聚集并凝聚形成胚状体。在短短的几天内，这些细胞便分化成许多不同的体细胞类型。

在未分化的阶段，XX 的胚胎干细胞有两条活化的 X 染色体。当细胞被诱导分化时，胚状体的大部分细胞都发生 X 染色体的随机失活。相反，在分化的 XY 胚胎干细胞中并不发生 *Xist* 的表达和 X 染色体的失活。因此，胚胎干细胞提供一个有价值的模式系统，它们可以通过基因打靶的方式来人工操纵基因，并通过这种方法对 X 染色体失活的不同阶段进行研究。一个重要的实验证明，即使在正在分化的 XY 胚胎干细胞中，敲除紧邻 *Xist* 下游 65kb 的片段可导致 *Xist* 等位基因的表达（因此而失活）（Clerc and Avner 1998）。这个发现提示我们，敲除的序列是假定的屏蔽因子结合所需要的。用更换靶位点的战略，敲除 65kb 内的更小区域来进一步描述这个区域作用的研究已经完成（图 17-6）。关键元件的确定和它们结合因子的鉴定是我们未来研究的主要目标。

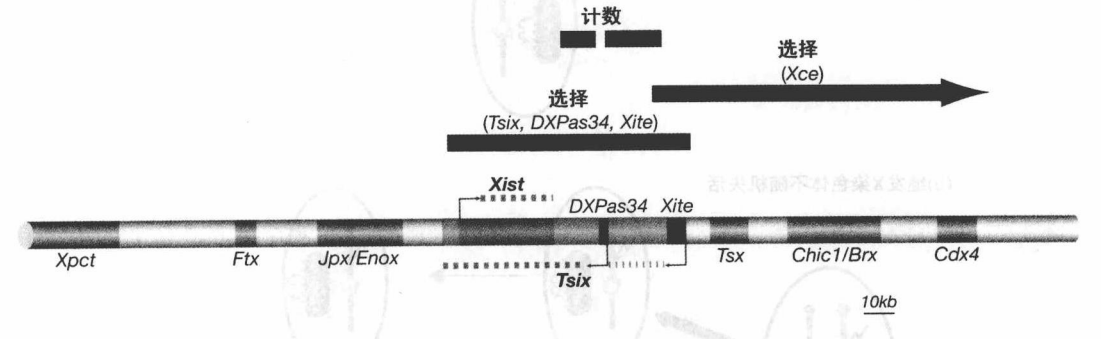



图 17-6　染色体失活中心区域的基因和调控元件

调控 X 染色体失活的关键区域用绿色表示。侧面的基因用灰色表示。箭头表示 *Xist*（正义）基因和 *Tsix*（反义）基因的启动子。各非编码 RNA 的延伸用绿色的虚线表示。调控元件控制 *Tsix*、*Xite* 和 *DXPas34* 的表达，用黑色表示。X 染色体选择、计数区域和基因座在上面表示。

## 3.4　X 染色体的随机失活的调控——选择

在特定的条件下，两条 X 染色体随机失活的概率可以不同。这可能是最初选择哪条 X 染色体失活时产生偏差的结果（初期 X 染色体不随机失活），也可能是对抑制维持特定 X 染色体活性的细胞选择的结果（续发的 X 染色体不随机失活）（图 17-7）。在初期 X 染色体非随机性失活中，这一选择受一些因素的影响。如顺式作用序列的差异，或杂合子动物中影响特定 X 染色体激活/失活可能性的突变。根据 Rastan 的屏蔽因子模型，这些差异可以通过改变屏蔽因子与特定等位基因结合的可能性来发挥它们的作用。

初期 X 染色体不随机失活的例子是小鼠中 X 染色体控制元件（*Xce*）。*Xce* 是一个典型的明确的基因座，已经发现这里不同的等位基因影响 XX 杂合子中 X 染色体被激活的可能（Cattanach and Isaacson，1967）。基因图位克隆把 *Xce* 定位在紧靠 *Xist* 的下游（图 17-6），因此正好在所谓 65kb 缺失后影响封闭因子结合的位点。

小鼠中影响选择的第二个元件是 *Tsix* 反义调控因子（Lee and Lu 1999）。这被认为是 *Tsix* 反义 RNA 介导的，在 X 染色体随机失活起始前通过 *Xist* 基因座转录（图 17-6）。在 X 染色体随机失活的 XX 细胞中，*Tsix* 启动子缺失的染色体更易失活。在控制 *Tsix* 表达水平的增强子元件（*Xist* 元件）突变的细胞中，可以观察到更多细微的差

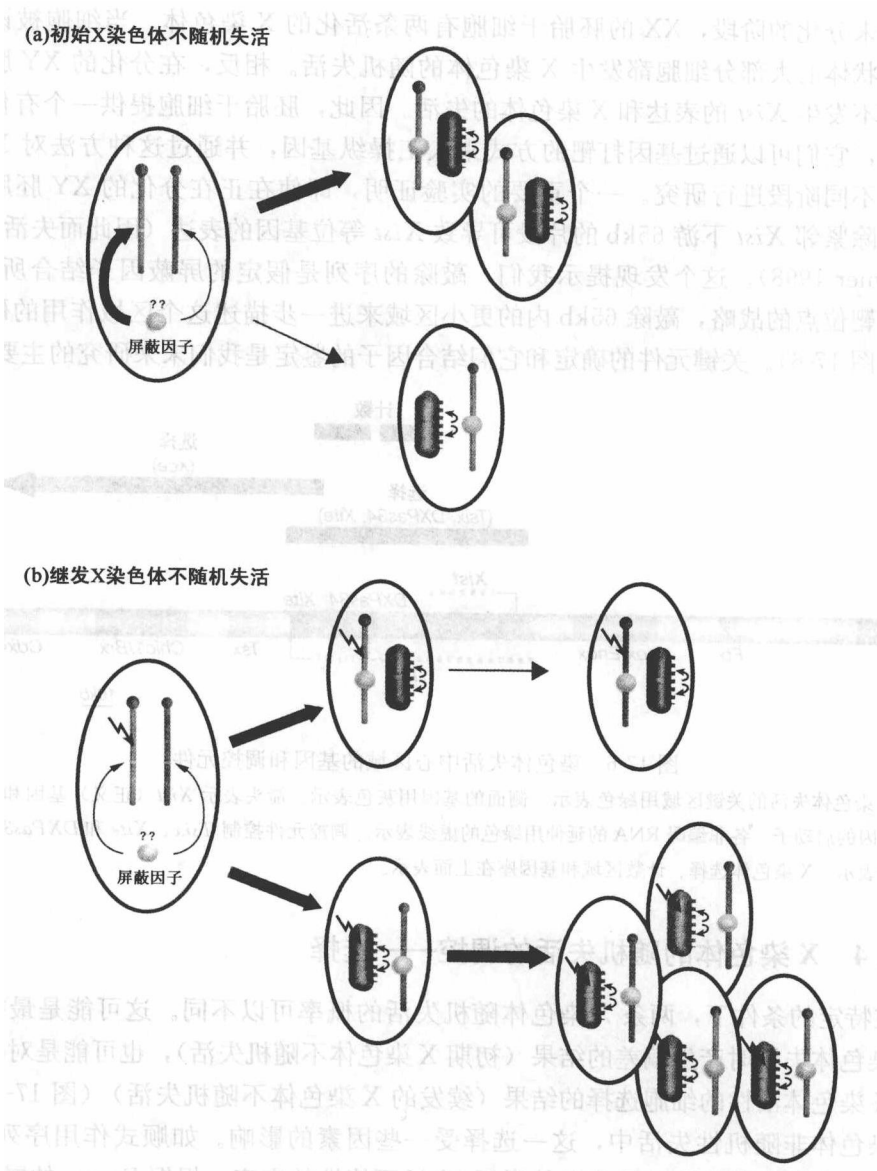



图 17-7 X 染色体不随机失活的模型

初始 X 染色体不随机失活指最初 X 染色体失活的选择具有倾向。理论上，在杂合的雌性中两个等位基因结合屏蔽因子有偏向的可能性。继发 X 染色体不随机失活的模型中，哪条染色体失活的选择是随机的，但细胞的选择导致两条 X 染色体中一条失活的细胞逐渐丢失。例如，在一条 X 染色体上有一个缺失突变位点的细胞，会使其他的位点，乃至整个染色体范围内的基因更易于丢失。

别。尽管 *Tsix* 的启动子所在的区域被定义为假定的封闭因子结合位点，但在 XY 胚胎干细胞中，定点缺失实验并不能通过缺失来激活 *Xist*。这就表示基因座并不相同，但缺失 *Tsix* 启动子区域可保持封闭因子结合位点的完整性。

在 X 染色体随机失活起始前，*Xist* 的转录伴随着 *Tsix* 的低水平转录，这表示双链

RNA（*Tsix*：*Xist* 杂链）可以介导选择作用。与此一致的是，通过 *Xist* 的启动子增加正义链的转录水平可挽救 *Tsix* 的抑制作用，使 XX 杂合子细胞中 X 染色体的等位基因更不容易被激活（Nesterova et al. 2003）。总之，在提到所有已知影响 X 染色体随机失活和印记失活的不同 *Xic* 元件，经常强调它们在影响计算和选择方面的功能，这在图 17-6 中显示。

### 3.5　早期胚胎基因失活的转换模式

小鼠早期胚胎是如何实现从印记调控模式到随机调控模式的转换过程的？直到最近认为这两种模式中 X 染色体失活的起始与细胞分化相关（Monk and Harper 1979）。因此，人们认为滋养外胚层（trophectoderm）和原始内胚层（primitive endoderm）细胞系中 Xp 的失活是为了适应胚泡期第一次分化时 *Xist* 上父本的印记。然而，产生完整胚胎的 ICM 细胞，认为第一次消除 *Xist* 印记，在原肠胚形成期分化成三种生殖细胞系时经历了 X 染色体的随机失活。但更近的数据说明在包括 ICM 前体细胞的所有细胞中，Xp 的失活发生在胚胎卵裂期细胞分化之前（Mak et al. 2004；Okamoto et al. 2004）。滋养外胚层（trophectoderm）和原始内胚层（primitive endodermt）细胞系中 X 染色体的印记失活是胚胎卵裂早期发生的 X 染色体失活模式的残留。因此，ICM 细胞必须启动一个程序来逆转 Xp 印记失活开始的波动，也为随后 X 染色体的随机失活提供条件（图 17-4）。失活的 Xp 再激活的基础仍不清楚，但这个过程可能涉及 ICM 特异性的抑制 Xp 中 *Xist* 表达的程序（见第 5 章）。

## 4. 失活状态的增殖和维持

### 4.1　*Xist* RNA、基因沉默和异染色质组装

*Xist* 基因到底什么作用？有力的证据证明 *Xist* 基因和其 RNA 产物是顺式启动 X 染色体失活的开关，也是沉默沿着染色体延伸的手段。有证据表明①*Xist* 只从 Xi 中表达；②在植入前胚胎 X 染色体失活时，*Xist* 的 RNA 水平迅速增加；③*Xist* 的表达水平在 X 染色体失活前上调，可能是其发生所完全必须；④*Xist* 的 RNA 与 Xi 共定位在间期核内，并沿着中期的两条染色体中的一条分布（参考图 17-3）；⑤插入在常染色体中含有 *Xist* 的转基因至少可诱导部分失活染色体的特性。过表达致使常染色体顺式包装成异染色质样的转录沉默染色体结构（Heard et al. 1999 及其参考文献）。这个发现说明 *Xist* 的 RNA 对启动异染色质的形成和转录沉默是必要而充分的。然而，X 染色体失活状态的维持不需要 *Xist* 的持续表达。例如，在人类和啮齿类（rodent）体细胞混合物中，*Xist* 在人类失活 X 染色体中表达丢失，但却保留在啮齿类，用来维持 X 染色体连锁基因的沉默（Brown and Willard 1994）。这个问题将在第 5 节中进一步讨论。

但必须注意的是，*Xist* 的 RNA 和 X 染色体失活的关系是选择性的。并没有发现他们沿着被持续激活的常染色质 PAR 以及组成性（着丝粒）异染色质。此外，中期染色体的分析显示有个带状的区域与基因丰富的 G 亮带相关（图 17-3）（Duthie et

al. 1999)，这个现象表明 *Xist* 的 RNA 仅包裹着染色体的特定区域（4.5 节将进一步讨论）。

*Xist* 的 RNA 引起染色质结构改变的详细机制以及引起相关基因沉默的详细信息现在仍不清楚。*Xist* RNA 分子的不同区域以及沿着 X 染色体分布并负责不同基因的沉默。在小鼠胚胎干细胞诱导 *Xist* 表达的实验中，可以检测特定缺失后 *Xist* 分子所具有的功能，实验表明，沉默是由于分子 5′端保守的重复序列引起，同时，染色体包裹是由散在分布于这个分子其余区域的序列所介导的。

## 4.2 失活 X 染色体的异染色质结构

从最早使用光学显微镜研究开始，人们已经认识到 X 染色体失活和异染色质有些共同的性质。如组成型异染色质在着丝粒及其周围时，Xi 仍然可见，它在整个分裂间期明显凝聚（如巴氏小体）并且其 DNA 在 S 期后复制。因此认为 Xi 包括功能型异染色质。然而，组成型异染色质 DNA 中有含量丰富的、特异的卫星重复序列，至少负责它们部分功能。尽管这些特异的重复元件有些精细的差别在失活过程中起一定的作用，但 X 染色体的重复 DNA 并不十分丰富（参考 4.5 节）。另外，失活 X 染色体通常被描述为“凝聚”，在 Xa 和 Xi 染色体上标记 X 染色体特异性探针、显微镜精细分析和三维重构的结果显示，其差别不仅仅在它们的形状上，而且在单位体积中染色体的数量也不同（Eils et al. 1996）。

通过间接的免疫荧光显微镜对组蛋白修饰的分布和中期染色体和间期核差别的研究，对 Xi 和组成性异染色质做进一步对比。在人类和小鼠的细胞中，失活 X 染色体的功能性异染色质中组蛋白 H4 的乙酰化消失（Jeppesen and Turner 1993），通过这种方式来重新组装组成型异染色质和着丝粒异染色质。这是首次证明失活的 X 染色体有被标记的特异的组蛋白修饰。接下来的实验是证明许多实验室观察到的现象，并进一步研究在间期和中期细胞中，组成型异染色质和功能型异染色质中所有四种核心组蛋白（H2A、H2B、H3 和 H4）的乙酰化形式（O'Neill et al. 2003；及其参考文献）。值得注意的是，在着丝粒异染色质和 Xi 中，在 H3K4 上存在的双甲基化和三甲基化缺失。像乙酰化一样，这通常被认为是转录激活或潜在激活的标记。

当考虑到这个标记的存在与转录沉默相关，但消失却与转录激活不相关时，这种情况变得更加复杂。例如，H3K9 的双甲基化和三甲基化通常是转录沉默基因的标记，荧光显微镜观察时，经常发现它在着丝粒异染色质中富集（Lachner et al. 2003）。但由于技术和生物学的原因，它们在 Xi 中富集的原因仍不清楚。因此，应进一步提高抗体特异性免疫荧光数据的重要性。例如，H3K9 和 K27 中均有部分 ARKS 四肽，两种抗体可能起交叉反应。这种交叉反应必须完全排除结果才能有效地说明问题。另外，使用的抗原和免疫过程均影响抗体的特异性。例如，由交联肽产生的抗 H3K9me3 的抗血清在异染色质区域中比在常染色质中组蛋白的修饰更稳定（Maison et al. 2002）。用这种方式产生的抗血清是研究异染色质非常有用，但不是定量比较异染色质和常染色质理想的手段。最终，间期的免疫荧光增强染色，仅能得到异染色质中高度致密的核体结果（Perche et al. 2000）。

在培养的人类细胞中，可以通过仔细分析沿着 Xi 组蛋白修饰的分布，进一步观察到这个系统的复杂性（Chadwick and Willard 2004）。H3K9me3 和 H3K27me3 均在特定的、不重叠的、穿过 Xi 的区域富集。因此，不像失去组蛋白的乙酰化，丰富的修饰是在 Xi 的一个区域，但这不是所有 Xi 的性质。有趣的是，在发现 H3K27me3 富集的区域同样发现 *Xist* RNA 和不同 macroH2A1.2 的富集（Constanzi and Pehrson 1998）。macroH2A1.2 和 Xi 的关系以及它在失活过程中的作用将在 4.4 节中继续讨论。相反，Xi 中 H3K9me3 富集的这些区域异染色质蛋白 HP1（已知与甲基化的 H3K9 相结合）和 H4K20me3（与着丝粒异染色质相关联的一个标志）的表达水平升高。重要的是，间期细胞巴氏小体的免疫染色显示相同的共染模式，这说明不同的区域是通过细胞周期来维持的。组蛋白修饰、组蛋白变体、非组蛋白质和 *Xist* RNA 共同出现并相互作用形成染色质，这些染色质具转录沉默的特性，在 S 后期复制，以“凝聚”的形式出现，并可能定位在核中，这些都是 Xi 的性质。但是，Xi 中特定染色体区域确切的功能意义仍然在被研究。一个重要的、也是让人迷惑的现象是，从一种人类细胞中到另一类人类细胞中观察到的区域出现的频率有很大的差别（Chadwick and Willard 2004）。这仅仅可证明在 X 染色体失活系统中已知的重复（如早期发现在成熟的细胞中，即使 *Xist* RNA 缺失后沉默仍能维持），但也让我们注意到，在培养的细胞系，尤其是免疫细胞系中发生的情况并不能完全正确地反映原代组织（primary tissues）中发生的情况，也当然不是发育的早期发生的情况。

组蛋白修饰与常染色质、组成性异染色质和功能性异染色质的关系总结在表 17-1 中。至今为止，仅仅在 Xi 中发现 H3K27 的甲基化，组蛋白 H2A K119 的泛素化的富集，但没有在组成型异染色质中发现（Plath et al. 2003；Silva et al. 2003；de Napoles et al. 2004；Smith et al. 2004）。

**表 17-1　组成型和功能型异染色质组蛋白修饰的特性**

| 组蛋白 | 修饰 | 数量(相对于常染色质) | |
|---|---|---|---|
| | | 组成型异染色质(中心的)? | 功能型异染色质(X 染色质失活) |
| H2A,H2B,H3,H4 | 乙酰化[a] | − | − |
| H2A | K119 素化 | ≈ | + |
| H3 | K4 双/三甲基化 | − | − |
| H3 | K9 双/三甲基化 | + | ≈[b] |
| H3 | K27 双/三甲基化 | ≈ | +[c] |
| H4 | K20 三甲基化 | + | ≈ |

注：表中指出，相对于常染色质，组成性异染色质和功能性异染色质中，组蛋白修饰被加强(+)或减少(−)。≈表示与常染色质相比，修饰水平没有检测到差异。(me)甲基化。

a 用于乙酰化的赖氨酸。b 在局部“热点(hot spot)”富集，但不是全部的。c 在一些类型细胞中短暂富集

在表 17-1 中列出的所有的修饰，都与失活 X 染色体的整个异染色质形成过程相关，这已经通过间期细胞中期染色体或巴氏小体的免疫荧光分析被识别。但是组蛋白修饰位置的改变可能也在 X 染色体失活不同阶段起着重要作用。这种变化可以通过高分辨率的显微镜识别，也可以通过 Xic 关键区域里或连接处染色体免疫共沉淀（ChIP）定位修饰的方法（第 10 章）来识别。例如，在未分化的胚胎干细胞中，*Xist* 基因 5′端延伸

的一大段超过 340kb 的区域具有 H3K9 高甲基化的性质。这个高甲基化的状态在细胞分化和 X 染色体失活过程中消失（Heard et al. 2001）。在雌性的胚胎干细胞中，同样的区域 H3K27 的甲基化水平也很高（Rougeulle et al. 2004），并且在这些区域内 H3 和 H4 位点的乙酰化水平也很高（O'Neill et al. 1999）。下一步研究将揭示 Xic 区域中什么程度的组蛋白修饰可启动早期 X 染色体失活过程，或下游染色体重塑过程元件的表达（可能有必要）。

组成型着丝粒异染色质中 DNA 甲基化状态富集，主要发生在 CpG 二聚体中 5′甲基胞嘧啶（见第 18 章）。这与它的低转录活性相一致。令人吃惊的是在 Xi 中 CpG 的甲基化整体上没有明显地高于其他的基因组。但是，与沉默基因相关的特异的 CpG 岛高甲基化。实验证据说明 DNA 甲基化对维持失活状态起着重要的作用。因此，小鼠缺少能甲基化以前未修饰 CpG 岛的酶（从头甲基化酶，Dnmta 和 Dnmtb），或缺少能维持之前修饰残基的修饰状态的酶（维持印基化酶，Dnmt1），均可以正常的方式启动或维持 X 染色体的随机失活过程（Sado et al. 2000 2004）。Xi 上的基因在 Xi 低甲基化的突变小鼠中，比在野生型正常甲基化水平的小鼠中更容易被再激活（Sado et al. 2000）。

## 4.3 X 染色体失活的组蛋白修饰酶学

迄今为止，还不清楚这些在 X 染色体失活（Xi）过程中负责核心组蛋白去乙酰化的酶和负责 H3K4 去甲基化的酶。因为去乙酰化酶抑制剂曲古菌素 A（trichostatin A，TSA）可以阻止乙酰化水平，至少可推迟乙酰化过程的发生。正在分化的雌性 ES 细胞中可出现去乙酰化的 X 染色体，因此可以认为 DHDAC 也参与这个过程（O'Neill et al. 1999）。但我们还不清楚这 11 种Ⅰ型和Ⅱ型 DHDAC 中哪种最负责这个过程（Ⅲ型 DHDAC 不受 TSA 的抑制）。我们也不知道负责把 H3K4 甲基化去除的机制。能把单甲基化和双甲基化的 H3K4 去甲基化的酶现在已经被找到，去三甲基化 H3K4 的酶仍在寻找。现在，我们还不能通过组蛋白替换或被动地通过 DNA 复制来把 H3K4 甲基化和组蛋白去乙酰化从 Xi 中去除。

现在我们已经知道了很多负责特定位点组蛋白修饰的酶，Ezh2/Enx1 负责 H3K27 的甲基化，它是果蝇 PcG 蛋白 E（Z）的同源物。E（Z）是一种组蛋白甲基化酶（histone methyltransferase，HKMT），在 PRC2 PcG 复合物中起作用（见第 11 章）。在 ES 细胞分化过程中，PRC2 招募 Xi 与 H3K27 甲基化的动力学相同。PcG 蛋白也参与 Xi 中 H2A K119 的泛素化。尤其是 Ring1A/Ring1B 蛋白，他们是 PRC1 /PcG 复合物的核心组成成分，和 E3 连接酶一样，负责 H2A 的泛素化。基因敲除 Ring1A 和 Ring1B 可使 Xi 上和基因组中 H2A 的泛素化丢失。

## 4.4 ES 细胞作为模式系统研究导致 X 染色体失活的有序过程

小鼠的 ES 细胞为研究 X 染色体失活的动力学提供了宝贵的模式系统。分化 1～2 天后的大部分细胞中可监测到 *Xist* 的 RNA 水平和对一条染色体的包裹程度的增加。有证据表明转录水平和转录后水平的调控均使 *Xist* 上调（Panning et al. 1997；Sheardown et al. 1997；Rougeulle et al. 2004）。然而，*Xist* 是从未分化雌性 ES 细胞中的两

条 X 染色体和雄性 ES 细胞中的一条染色体转录而来，但其 RNA 产物迅速降解，仅能检测到靠近 *Xist* 基因座的一小部分。有证据表明，在分化过程中雌性两条 X 染色体的一条上的 *Xist* RNA 是稳定的（Panning et al. 1997；Sheardown et al. 1997）选择 RNA 稳定步骤的基础机制仍不清楚，因为这有助于整体增加 *Xist* RNA 水平和染色体顺式包装程度。

已经发现在正在分化的 XX ES 细胞中，一些 X 染色体失活步骤与 *Xist* RNA 的开始积累同时发生。这诱导招募 PcG 蛋白，并与 H3K27 甲基化、H2A 的单泛素化、H3K9 的去乙酰化、H3K4 甲基化的丢失相关（Heard et al. 2001；Silva et al. 2003；de Napoles et al. 2004；Rougeulle et al. 2004）。整个组蛋白的去乙酰化是相对较晚的事件，它发生在大多数细胞 3～5 天时，可能参与失活状态的维持或稳定，而不是起始（Keohane et al. 1996）。这也解释了个别经历失活的基因乙酰化的形式在启动子上，这些可通过整个染色体或大量区域的免疫荧光分析测定得出。最初通过 ChIP 的研究说明情况确实如此，但还有必要通过大量基因的进一步实验（O'Neill et al. 2003）。

变体组蛋白 macroH2A1. 2 在 Xi 的积累发生在 XX ES 细胞分化的较晚期（Mermoud et al. 1999）。这种变体组蛋白整个分子在 C 端尾巴上有超过 200 个多余的氨基酸和一些被置换的氨基酸。有趣的是，体细胞中 *Xist* RNA 的表达需要 macroH2A 保留在 Xi 上，但在早期分化阶段 Xi 并足以不招募 macroH2A（Mermoud et al. 1999；Wutz et al. 2002）。

在 ES 细胞中，在 Xi 上选择 DNA 甲基化是一个更晚发生的过程。众所周知，在成年细胞中的 Xi 上 CpG 高甲基化，但在雌性的 ES 细胞中，CpG 直到分化后 14～21 天才被甲基化（Keohane et al. 1996），这与自身胚胎发育的结果（Lock et al. 1987）以及 DNA 甲基化负责稳定和保持、而不是起始和延伸失活状态是相一致的。

因此，在发育过程中，具有协同和精细调控功能的序列的出现可使 Xi 上染色体改变（图 17-8）。值得注意的是，一些这样的改变，如组蛋白去乙酰化和 DNA 甲基化是在细胞起始下游分化的各种不同途径后发生的。看起来，这个过程负责 X 染色体失活的过程独立于其他细胞分化的过程外完成。但是需要注意到，X 染色体随机失活的一些方面只有在细胞分化开始后才能进行。例如，在未分化 ES 细胞中 *Xist* 转基因的表达启动与异染色质相关的一些组蛋白修饰过程的发生，也是 S 期后复制过程的过渡（Wutz and Jaenisch 2000），但没有检测到掺入的 macroH2A。只有细胞诱导分化后，macroH2A 和 *Xist* RNA 才能在含有 *Xist* 转基因的染色体上共定位（Rasmussen et al. 2001）。macroH2A 与 *Xist* 包裹的染色体的联系，在于不断产生 *Xist* RNA（Csankovszki et al. 1999）而不需要转录沉默。因为在沉默所需区域突变的 *Xist* RNA 上，仍然发现 *Xist* RNA 包裹染色体（Wutz et al. 2002）。因此，X 染色体失活过程可以看做是一系列平行过程的结果，仅有一部分是相互依赖的。

特有的酶复合物和组蛋白修饰可能在 X 染色体失活过程的特定阶段具有重要功能，强调这一点是非常重要的，但后来，可能在 DNA 甲基化引起更多恒定的沉默机制发生后，也可能变得不太重要。例如，PRC2 催化的 H3K27 甲基化是早期个体发育中 X 染色体成功失活所必需的，但似乎后来消失了。

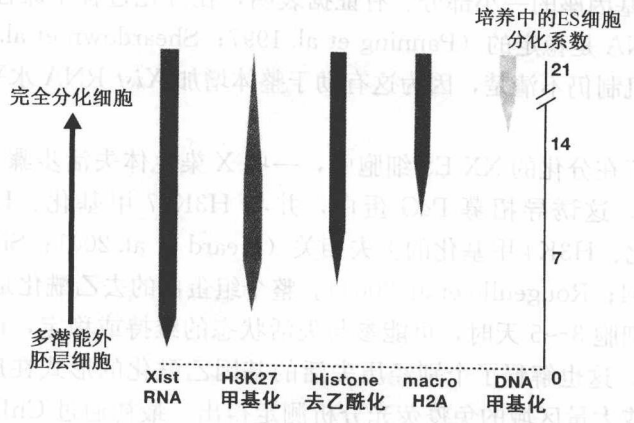



图 17-8 通过分化 Xi 上不同层次表观遗传沉默的积累

在发育的胚胎和培养的 ES 细胞中，与转录沉默相关的五种不同的表观遗传改变在发育和分化的不同阶段在失活的 X 染色体上的发生不同。在一些类型细胞中，H3K27 的甲基化是瞬时的，其仅在分化的早期阶段被发现，在成熟的细胞中不存在。

应该强调，在植入前，胚胎中 X 染色体印记失活完成的过程中，不按次序的事件可能发生。主要是 H3K27me3 富集的甲基化状态只有到了 16 细胞期才被检测到，远远迟于 *Xist* 开始表达时（2～4 细胞期）（Mak et al. 2004；Okamoto et al. 2004）。这表明需要特异的发育调控辅助因子来招募 PRC2 PcG 复合物到 Xi 上。

## 4.5 沉默染色质的蔓延

Xic 是 X 染色体失活过程所必需的，有人认为沉默的蔓延是从 Xic 开始的，邻近的基因比稍远的基因沉默的早。一种解释的理论是 *Xist* RNA 从合成位点开始，沿着染色体逐渐的蔓延。这与 *Xist* 转基因插入常染色体中，可使常染色体被 *Xist* RNA 包裹，导致基因沉默和染色质改变的实验结果是相一致的。

失活状态从 Xic 蔓延的研究已经扩展到自然发生的 X 染色体、常染色体异位。的确，这种异位对说明 Xic 的存在是非常关键的（Rastan 1983）。具有功能性异染色质标记的染色质（如组蛋白乙酰化的丢失、复制晚和转录沉默）已经证明是从 Xi 到混合染色体的常染色体部分蔓延的（White et al. 1998；Duthie et al. 1999；Sharp et al. 2002）。这就形成一个重要的原则，功能性异染色质并不是专指 X 染色体，这与常染色体上 *Xist* 转基因的结果是一致的。但在移位和有限的范围内，沉默染色质沿着常染色质臂的蔓延是不同的。

有两种模型可解释沉默只有限地蔓延至顺式连接的常染色体。①常染色体可以耐受失活 X 染色体上 *Xist* RNA 的起始蔓延和相关基因的沉默。②起始蔓延到常染色体是有效的，但在个体发育过程中不能很好的维持沉默状态，这叫做蔓延和退缩（spread and retreat)”。迄今的证据支持蔓延和退缩的模型，但这两种模型并不相互排斥，还需要进一步研究。应该注意是发生在常染色体上蔓延沉默的细胞，因为常染色体关键基因表达

的丢失，可以在早期发育过程中来选择抑制。

在一些 X 染色体上，常染色体转录的蔓延是不连续的，它跳过一些特定的区域，离开转录活跃区，被沉默区包围的常染色体区和异染色质区（Sharp et al. 2002）。组成性异染色质蔓延到连接常染色质的连接区，这些区域在果蝇中类似，可引起位置效应花斑（见第 5 章）。这些现象在以沉默的蔓延和退缩为基础的早期细胞选择机制中比在稳定沉默的继续延伸时更容易停止。

有人认为，沿着 X 染色体分布着元件，这些元件叫做“站点（way station）”，负责异染色质的组装位点，因此可增强失活 X 染色体的蔓延和维持（Gartler and Riggs 1983）。这些元件必须不普遍，也不是有规则地分布在常染色体上。有人认为有一个通常弥散分布的重复序列家族，散布的长重复序列（long interspersed repeat，LINE）是这些“站点”元件好的候选者（Lyon 1998 2003）。这些重复序列在人和小鼠的基因组中相同，在 X 染色体的分布尤为广泛。此外，LINE 元件都在人和小鼠的基因组中聚集浓缩、基因较少的 G 带区，这说明它们可能在某种程度上青睐转录沉默相关的染色体构象。最近，人类 X 染色体上 DNA 的测序已经完成，揭示了广泛认为是“站点”的 LINE 元件的分布，但 LINE 元件的真正角色还未证明。

## 4.6　X 染色质逃避失活

以前讨论过，一些基因逃避失活。就机制而言，X 染色体失活的逃避是由于连接在这些基因上的“站点”很少，或是由于阻碍 *Xist* RNA 或其他沉默成分蔓延的边界元件的出现。有证据表明，在早期发育中，逃避失活的基因至少在某种程度上是沉默的。例如，小鼠中 *Smcx* 基因可通过改变发育阶段和组织类型来逃避沉默（Sheardown et al. 1996）。因此，X 染色体上沉默常染色体基因或异位的背景下，这与前面所讨论的蔓延和退缩理论是一致的。

## 4.7　袋类哺乳动物中的 X 染色体失活

真哺乳亚纲（Eutherian）的哺乳动物与袋类（Marsupial）哺乳动物的分化是在大约 1.3 亿年前。和真哺乳亚纲动物一样，袋类动物也是 XY（雄性）和 XX（雌性）的性别决定系统，采用雌性两条 X 染色体中一条失活的剂量补偿机制。正如前面提到的，袋类动物失活的 X 染色体总是父本来源的染色体（Xp）。与真哺乳亚纲动物差别的进一步对比，在不同组织中，个体基因座的沉默程度不同。也有些证据证明，随着发育和个体发生，沉默的易变性程度增加。这是真哺乳亚纲动物中常染色质基因座沉默的蔓延和退却模型的又一暗示。有趣的是，X 染色体上 LINE 元件的富集是在真哺乳亚纲动物和后兽亚纲动物分化时才开始发生的，因此，袋类动物 X 染色体上沉默的易变性也可能与“站点”的观点有关。此外，袋类动物中 CpG 岛中甲基化的缺少与 X 连锁基因的关系加起来有助于 Xi 的不稳定性。

现在，对袋类动物中 X 染色体失活的分子机制了解很少，也没有找到 *XIST* 在袋类动物中的同源物。当然，虽然现在没有发现，但 *XIST* 同源物不存在并不能排除能够起始 X 染色体失活的 RNA 同源物的不存在。在真哺乳亚纲动物和袋类动物失活的 X

染色体中，都明确的具有两种性质，它们都在S期后复制，并都具有组蛋白H4低水平乙酰化的标志（Wakefield et al. 1997）。

考虑到在一些生物中，包括其他的哺乳动物，破坏剂量补偿作用可使雄性致死，袋类动物能够耐受不完全的剂量补偿作用的原因的却让人迷惑。X染色体上袋类动物所带的基因比真哺乳亚纲动物少是一种可能的解释（Graves 1996；Graves and Shetty 2001）。袋类动物中X染色体上仅仅出现在真哺乳亚纲动物X染色体的长臂上出现的基因。在真哺乳亚纲动物X染色体的短臂上出现的基因分布在袋类动物的常染色体上。因此，袋类动物X染色体上一大部分都是基因较少和组成性异染色质。可能这种基因数量的减少可使一些类型细胞能够耐受剂量补偿机制的减轻，但不足以生物全部免除剂量补偿作用。另外，袋类动物在早期发育过程中可更有效地沉默Xi，就像是印记基因座中的过程一样，这时剂量的差异可能最关键。这两种解释也可能都有一定的作用。

## 5. X染色体的再激活和重编程

### 5.1 体细胞中X染色体失活的稳定性

多层次的表观遗传修饰使X染色体失活沉默，因此，通常抑制状态是很稳定的。在早期已经得知，DNA甲基化的抑制剂5-氮胞苷（5-azacytidine，5azaC），在XX细胞中可逆转X染色体沉默的状态（Mohandas et al. 1981）。个别基因的散在再激活可在较低的频率下发生。但是，特定基因再激活后对细胞系的分析显示其他基因通常仍是失活的。与估计一致的是，在整个染色体水平上，用TSA抑制Ⅰ型组蛋白去乙酰化也可以得到相同的数据（Ccankovszki et al. 2001）。

在小鼠衰老过程中，可以观察到个别X染色体连锁基因的散在再激活。在袋类动物个体发生过程中失活的X染色体不那么稳定并逐渐发生，这种情况甚至有优先发生的倾向。

*Xist* RNA的作用是什么呢？用人类细胞系得到的数据显示，在X染色体上缺失包围*XIST*的区域后并没检测到X染色体的再活化（Brown and Willard 1994）。这个结果在小鼠成纤维细胞中，条件性去除*Xist*等位基因时被再次证明（Csankovszki et al. 1999）。在这里展示了*Xist*的缺失可导致组蛋白变体macroH2A从Xi上的离开。更多最近的研究说明，在条件性去除*Xist*的成纤维细胞中，被PcG复合物催化的组蛋白修饰、H3 K27三甲基化、H2A的泛素化也丢失（Plath et al. 2004）。个别基因很少检测到散在的再激活，在用5azaC或TSA处理后这种情况更不常见。但没有整个染色体范围内明显再激活发生。因此，移去多个表观遗传沉默的标记仍不足以逆转染色体范围内的沉默。

在体细胞中，Xi沉默过多的机制叫做“纽带机制（belts and braces）”。这也暗示个体各个水平的表观遗传沉默既是自我维持的机制，又是通过相关的正反馈机制维持的过程。在其他的表观遗传系统中，也观察到类似的正反馈机制，如臂间异染色质中组蛋白甲基化和DNA甲基化的维持。

## 5.2　正常发育中 X 染色体的再激活

虽然在体细胞中失活的 X 染色体非常稳定，但在正常发育过程的一些情况下可使整条染色体被再激活。最好的例子是在发育中的原生殖细胞内（primordial germ cell，PGC），失活的 X 染色体可以被再激活。在小鼠中，PGC 细胞特化发生在发育的 7～8 天，即原肠胚刚刚形成后。在这个时期，胚胎的细胞已经经过了 X 染色体的随机失活过程。后来，发育中的 PGC 沿着胚胎的后肠区延伸，到达生殖脊——产生性腺的结构。也就是在这个时候，XX PGC 再激活它的 Xi（Monk and McLaren 1981）。这个过程与一个很普遍的表观遗传再编程过程一致，其中包括去除父本的印记和基因组范围的甲基化（详见第 20 章）。

PGC 中 X 染色体再激活是一个多层异染色质结构逆转的特定机制。已经发现 *Xist* RNA 表达的衰减与 X 染色体再激活过程是一致的，但如果 XX 体细胞中沉默是不依赖的 *Xist* 的，这就不一定是成因了。PGC 可能没有建立所有与沉默相关的标记，因此更容易被再激活。在小鼠中，发育的 PGC 的 Xi 上，CpG 岛并没有发生甲基化，这与前述也是一致的（Grant et al. 1992）。

X 染色体再激活过程第二个例子是在 3.5 节讨论过的，在胚胎囊泡期 ICM 定位过程中印记失活的 Xp 的再激活，这与更广泛的基因组内再编程过程有关。这个再激活的过程与 *Xist* RNA 表达的降低和一些沉默相关的表观遗传修饰标记的丢失有关。

## 5.3　重编程实验中 X 染色体的再激活

在特定的实验条件下可以观察到 X 染色体的再激活。把体细胞的核转移至未受精的卵母细胞中（Eggan et al. 2000），或是把体细胞和全能细胞，如 ES、胚胎生殖细胞（embryonic germ，EG）、胚胎癌性细胞（embryonal carcinoma，EC）融合后也可以发生这种情况（Tada et al. 2001）。

核转移胚胎有些特别有趣的例子。小鼠的实验研究说明在分裂期的核转移胚胎中，Xi 上的标记基因被迅速地再激活。尽管如此，克隆胚胎在胎儿阶段细胞核保留了一些关于哪条 X 是失活的记忆，供体细胞 Xi 也就是在胎盘滋养外胚层细胞的 Xi。相反，胚胎固有的细胞表现出 X 染色体的随机失活（图 17-9）。发育中 ICM 发生的 X 染色体再激活和重编程过程可能给胚胎一个表观遗传重新设定的第二次机会。

XX 体细胞和多能性胚胎细胞的融合细胞中，X 染色体的再激活还没有很好地研究。让体细胞的基因组接触 ES、EG、EC 细胞中的因子也可能发生这种情况。再激活过程的发生非常快，大约在融合后的第 5 天，经长时间培养后 *Xist* 的高表达状态在融合细胞中消失。

## 5.4　诱导 *Xist* 转基因表达的启示

在 ES 细胞中，一系列诱导 *Xist* 转基因表达的实验已经增加了我们对 X 染色体失活的稳定性和逆转性的理解。首先，已经证明在未分化的 ES 细胞中和细胞分化的早期，*Xist* RNA 可以建立 X 染色体的失活，后来的细胞却失去此功能（Wutz and Jae-

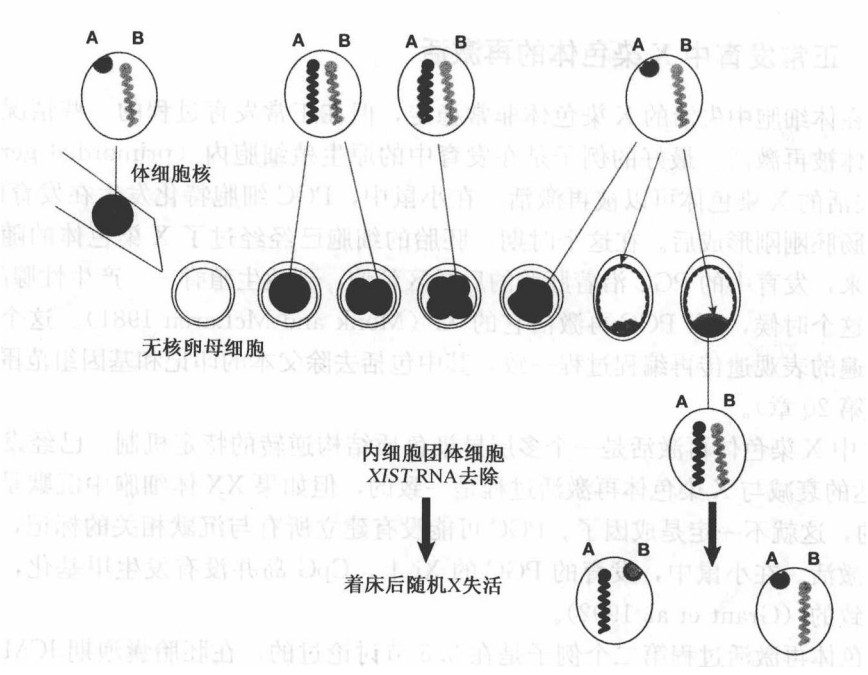



图 17-9 克隆的小鼠胚胎 X 染色体失活的调控

这个图展示了一个含有被 *Xist* RNA 包被的失活的 X 染色体的 XX 供体细胞。在该模型中，供体核基因的转录（包括 *Xist* RNA 在内）在细胞期之前都被卵母细胞因子所抑制，导致 X 染色体失活。在细胞期，供体细胞的 *Xist* 表达重新开始，而宿主细胞失活 X 染色体上的 *Xist* 也再次表达。这可能会在 *Xist* 启动子上加一个标记（如甲基化）。这样模式仅在 TE 和 DE 来源的细胞中发生，而多能性外胚层细胞因 *Xist* 表达再次沉默而没有。在内细胞团中，表观遗传标记控制宿主 *Xist* 的表达，使胚胎中接下来的 X 染色体随机失活过程正确进行。

nisch 2000)，这被称为“机遇窗口（window of opportunity)”。受 *Xist* RNA 影响的细胞功能可以推广到与 X 染色体失活的联系。因此，当在 ES 细胞或在早期分化阶段中转基因被关闭时可以逆转沉默状态，但在分化后期和体细胞中却不存在这种情况。

在 X 染色体再激活和重编程中，诱导转基因的数据强调在特定的细胞环境中，也就是未分化的 ES 细胞中，当 *Xist* RNA 的表达减少时，X 染色体再激活的过程就会发生。

如果我们考虑到那些已经发生 X 染色体再激活的细胞（如 PGC、ICM、EG 和 EC 细胞）在潜能性和可塑性方面都类似于 ES 细胞，那么在所有情况中，*Xist* 表达的减少可能是 X 染色体再激活的基础。

## 6. 总结和展望

在最近几年里，人们对 X 染色体失活分子机制的了解已经有了很大进步。目前为止，这些进步主要是由于表观遗传学相关领域以及其他受影响领域的进步。后一种情况的例子是，一些印记基因簇受顺式作用的非编码 RNA 的调控，如 *Xist* 对 X 染色体的调控（见第 19 章）。有理由认为这种互补性的进展将继续。

但是，现在仍有一些没有解决的问题。值得注意的是，尽管经过 40 年的研究，我们仍不明白，甚至只是大概了解“计数”和“选择”的机制。现在屏蔽因子假说已经有 20 年了，虽然它给我们提供了一些概念性指导，但对可能存在的本质，我们仍不了解。在对调控“计数”的特定顺式作用序列和反式作用因子方面的研究也取得了一些进步，进一步的研究也是令人兴奋的挑战。同样，尽管我们现在了解一些染色体修饰复合物：包括维持染色体失活状态的，如聚梳类复合物（Polycomb-group complexes）；建立基因组范围内沉默信号的，由 *Xist* RNA 引起的，但对其本质我们仍不了解。可能想到的是，我们需要了解在失活 X 染色体上，组蛋白去乙酰化和去甲基化的机制。我们需要了解的其他重要问题还有沉默是如何通过染色体蔓延的，如果有“站点”的话（也许是 LINE 元件），它在这个过程中和沉默状态的稳定或维持起到什么作用？这可能涉及一个令人好奇的问题：X 染色体失活状态是如何在一些类型细胞和发育阶段中逆转，而在其他情况下却不能逆转。后者则关系到一个更广泛和至关重要的话题：发育过程中基因组的可塑性和重编程过程。

吕雯雯　译
王大亮　校
孙方霖　审校

## 参 考 文 献

Beutler E., Yeh M. and Fairbanks V.F. 1962. The normal human female as a mosaic of X-chromosome activity: Studies using the gene for glucose-6-phosphate dehydrogenase deficiency as a marker. *Proc. Natl. Acad. Sci.* **48:** 9–16.

Brockdorff N., Ashworth A., Kay G.F., McCabe V.M., Norris D.P., Cooper P.J., Swift S., and Rastan S. 1992. The product of the mouse *Xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71:** 515–526.

Brown C.J. and Willard H.F. 1994. The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature* **368:** 154–156.

Brown C.J., Ballabio A., Rupert J.L., Lafreniere R.G., Grompe M., Tonlorenzi R., and Willard H.F. 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349:** 38–44.

Brown C.J., Hendrich B.D., Rupert J.L., Lafreniere R.G., Xing Y., Lawrence J., and Willard H.F. 1992. The human *XIST* gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71:** 527–542.

Carrel L. and Willard H.F. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* **434:** 400–404.

Cattanach B.M., and Isaacson J.H. 1967. Controlling elements in the mouse X chromosome. *Genetics* **57:** 331–346.

Chadwick B.P. and Willard H.F. 2004. Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proc. Natl. Acad. Sci.* **101:** 17450–17455.

Clerc P. and Avner P. 1998. Role of the region 3′ to *Xist* exon 6 in the counting process of X-chromosome inactivation. *Nat. Genet.* **19:** 249–253.

Charlesworth B. 1996. The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol.* **6:** 149–162.

Costanzi C. and Pehrson J.R. 1998. Histone macroH2A is concentrated in the inactive X chromosome of female mammals. *Nature* **393:** 599–601.

Csankovszki G., Nagy A., and Jaenisch R. 2001. Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *J. Cell Biol.* **153:** 773–784.

Csankovszki G., Panning B., Bates B., Pehrson J.R., and Jaenisch R. 1999. Conditional deletion of *Xist* disrupts histone macroH2A localization but not maintenance of X inactivation. *Nat. Genet.* **22:** 323–324.

Davidson R.G., Nitowsky H.M., and Childs B. 1963. Demonstration of two populations of cells in the human female heterozygous for glucose-6-phosphate dehydrogenase variants. *Proc. Natl. Acad. Sci.* **50:** 481–485.

de Napoles M., Mermoud J.E., Wakao R., Tang Y.A., Endoh M., Appanah R. Nesterova T.B., Silva J., Otte A.P., Vidal M., et al. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell* **7:** 663–676.

Duthie S.M., Nesterova T.B., Formstone E.J., Keohane A.M., Turner B.M., Zakian S.M., and Brockdorff N. 1999. *Xist* RNA exhibits a banded localization on the inactive X chromosome and is excluded from autosomal material in *cis*. *Hum. Mol. Genet.* **8:** 195–204.

Eggan K., Akutsu H., Hochedlinger K., Rideout W., III, Yanagimachi R., and Jaenisch R. 2000. X-chromosome inactivation in cloned mouse embryos. *Science* **290:** 1578–1581.

Eils R., Dietzel S., Bertin E., Schrock E., and Speicher M.R. 1996. Three dimensional reconstruction of painted human interphase chromosomes: Active and inactive X chromosome territories have similar volumes but differ in shape and surface structure. *J. Cell. Biol.* **135:** 1427–1440.

Epstein C.J., Smith S., Travis B., and Tucker G. 1978. Both X chromosomes function before visible X-chromosome inactivation in female mouse embryos. *Nature* **274:** 500–502.

Gartler S.M. and Riggs A.D. 1983. Mammalian X-chromosome inactivation. *Annu. Rev. Genet.* **17:** 155–190.

Grant M., Zuccotti M., and Monk M. 1992. Methylation of CpG sites of two X-linked genes coincides with X-inactivation in the female mouse embryo but not in the germ line. *Nat. Genet.* **2:** 161–166.

Graves J.A. 1996. Mammals that break the rules: Genetics of marsupials and monotremes. *Annu. Rev. Genet.* **30:** 233–260.

Heard E., Mongelard F., Arnauld D., Chureau C., Vourch C., and Avner P. 1999. Human *XIST* yeast artificial chromosome transgenes show partial X inactivation center function in mouse embryonic stem cells. *Proc. Natl. Acad. Sci.* **96:** 6841–6846.

Heard E., Rougeulle C., Arnaud D., Avner P., Allis C.D., and Spector D.L. 2001. Methylation of histone H3 at lys-9 is an early mark on the X chromosome during X inactivation. *Cell* **107:** 727–738.

Huynh K.D. and Lee J.T. 2003. Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426:** 857–862.

Jeppesen P. and Turner B.M. 1993. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation: A cytogenetic marker for gene expression. *Cell* **74:** 281–289.

Keohane A.M., Belyaev N.D., Lavender J.S., O'Neill L.P., and Turner B.M. 1996. X-inactivation and H4 acetylation in embryonic stem cells. *Dev. Biol.* **180:** 618–630.

Lachner M., O'Sullivan R.J., and Jenuwein T. 2003. An epigenetic road map for histone lysine methylation. *J. Cell Sci.* **116:** 2117–2124.

Lee J.T. 2000. Disruption of imprinted X inactivation by parent-of-origin effects at *Tsix*. *Cell* **103:** 17–27.

Lee J.T. and Lu N.F. 1999. Targeted mutagenesis of *Tsix* leads to nonrandom X inactivation. *Cell* **99:** 47–57.

Lee J.T., Davidow L.S., and Warshawsky D. 1999. *Tsix*, a gene antisense to *Xist* at the X-inactivation centre. *Nat. Genet.* **21:** 400–404.

Lee J.T., Strauss W.M., Dausman J.A., and Jaenisch R. 1996. A 450 kb transgene displays properties of the mammalian X-inactivation center. *Cell* **86:** 83–94.

Lock L.F., Takagi N., and Martin G.R. 1987. Methylation of the *Hprt* gene on the inactive X occurs after chromosome inactivation. *Cell* **48:** 39–46.

Lyon M.F. 1961. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature* **190:** 372–373.

———. 1998. X-chromosome inactivation: A repeat hypothesis. *Cytogenet. Cell Genet.* **80:** 133–137.

———. 2003. The Lyon and the LINE hypothesis. *Semin. Cell Dev. Biol.* **14:** 313–318.

Maison C., Bailly D., Peters A.H., Quivy J.P., Roche D., Taddei A., Lachner M., Jenuwein T., and Almouzni G. 2002. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* **30:** 329–334.

Mak W., Nesterova T.B., de Napoles M., Appanah R., Yamanaka S., Otte A.P., and Brockdorff N. 2004. Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303:** 666–669.

Marin I., Siegal M.L., and Baker B.S. 2000. The evolution of dosage-compensation mechanisms. *BioEssays* **22:** 1106–1114.

Marshall Graves J.A. and Shetty S. 2001. Sex from W to Z: Evolution of vertebrate sex chromosomes and sex determining genes. *J. Exp. Zool.* **290:** 449–462.

McElreavey K., Barbaux S., Ion A., and Fellous M. 1995. The genetic basis of murine and human sex determination: A review. *Heredity* **75:** 599–611.

Mermoud J.E., Costanzi C., Pehrson J.R., and Brockdorff N. 1999. Histone MacroH2A1.2 relocates to the inactive X chromosome after initiation and propagation of X-inactivation. *J. Cell Biol.* **147:** 1399–1408.

Mohandas T., Sparkes R.S., and Shapiro L.J. 1981. Reactivation of an inactive human X chromosome: Evidence for X inactivation by DNA methylation. *Science* **211:** 393–396.

Monk M. and Harper M.I. 1979. Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. *Nature* **281:** 311–313.

Monk M. and McLaren A. 1981. X-chromosome activity in foetal germ cells of the mouse. *J. Embryol. Exp. Morphol.* **63:** 75–84.

Nesterova T.B., Johnston C.M., Appanah R., Newall A.E.T., Godwin J., Alexiou M., and Brockdorff N. 2003. Skewing X chromosome choice by modulating sense transcription across the *Xist* locus. *Genes Dev.* **17:** 2177–2190.

Norris D.P., Patel D., Kay G.F., Penny G.D., Brockdorff N., Sheardown S.A., and Rastan S. 1994. Evidence that random and imprinted *Xist* expression is controlled by preemptive methylation. *Cell* **77:** 41–51.

Ohno S. 1967. *Sex chromosomes and sex-linked genes.* Springer-Verlag, Berlin pp. 1–140.

Okamoto I., Otte A.P., Allis C.D., Reinberg D., and Heard E. 2004. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303:** 644–649.

Okamoto I., Arnaud D., Le Baccon P., Otte A.P., Disteche C.M., Avner P., and Heard E. 2005. Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* **438:** 297–298.

O'Neill L.P., Randall T.E., Lavender J., Spotswood H.T., Lee J.T., and Turner B.M. 2003. X-linked genes in female embryonic stem cells carry an epigenetic mark prior to the onset of X inactivation. *Hum. Mol. Genet.* **12:** 1783–1790.

O'Neill L.P., Keohane A.M., Lavender J.S., McCabe V., Heard E., Avner P., Brockdorff N., and Turner B.M. 1999. A developmental switch in H4 acetylation upstream of *Xist* plays a role in X chromosome inactivation. *EMBO J.* **18:** 2897–2907.

Panning B., Dausman J. and Jaenisch R. 1997. X chromosome inactivation is mediated by *Xist* RNA stabilization. *Cell* **90:** 907–916.

Penny G.D., Kay G.F., Sheardown S.A., Rastan S., and Brockdorff N. 1996. Requirement for *Xist* in X chromosome inactivation. *Nature* **379:** 131–137.

Perche P.Y., Vourc'h C., Konecny L., Souchier C., Robert-Nicoud M., Dimitrov S., and Khochbin S. 2000. Higher concentrations of histone macroH2A in the Barr body are correlated with higher nucleosome density. *Curr Biol.* **10:** 1531-1534.

Plath K., Talbot D., Hamer K.M., Otte A.P., Yang T.P., Jaenisch R., and Panning B. 2004. Developmentally regulated alterations in Polycomb repressive complex 1 proteins on the inactive X chromosome. *J. Cell Biol.* **167:** 1025–1035.

Plath K., Fang J., Mlynarczyk-Evans S.K., Cao R., Worringer K.A., Wang H., de la Cruz C.C., Otte A.P., Panning B., and Zhang Y. 2003. Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300:** 131–135.

Rasmussen T.P., Wutz A.P., Pehrson J.R., and Jaenisch R.R. 2001. Expression of *Xist* RNA is sufficient to initiate macrochromatin body formation. *Chromosoma* **110:** 411–420.

Rastan S. 1983. Non-random inactivation in mouse X-autosome translocation embryos—Location of the inactivation centre. *J. Embryol. Exp. Morphol.* **78:** 1–22.

Rougeulle C., Chaumeil J., Sarma K., Allis C.D., Reinberg D., Avner P., and Heard E. 2004. Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol. Cell. Biol.* **24:** 5475–5484.

Sado T., Okano M., Li E., and Sasaki H. 2004. De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development* **131:** 975–982.

Sado T., Fenner M.H., Tan S.S., Tam P., Shioda T., and Li E. 2000. X inactivation in the mouse embryo deficient for *Dnmt1*: Distinct effect of hypomethylation on imprinted and random X inactivation. *Dev. Biol.* **225:** 294–303.

Sharman G.B. 1971. Late DNA replication in the paternally derived X chromosome of female Kangaroos. *Nature* **230:** 231–232.

Sharp A.J., Spotswood H.T., Robinson D.O., Turner B.M., and Jacobs P.A. 2002. Molecular and cytogenetic analysis of the spreading of X inactivation in X;autosome translocations. *Hum. Mol. Genet.* **11:** 3145–3156.

Sheardown S., Norris D., Fisher A., and Brockdorff N. 1996. The mouse *Smcx* gene exhibits developmental and tissue specific variation in degreee of escape from X inactivation. *Hum. Mol. Genet.* **5:** 1355–1360.

Sheardown S.A., Duthie S.M., Johnston C.M., Newall A.E.T., Formstone E.J., Arkell R.M., Nesterova T.B., Alghisi G-C., Rastan S., and Brockdorff N. 1997. Stabilization of *Xist* RNA mediates initiation of X chromosome inactivation. *Cell* **91:** 99–107.

Silva J., Mak W., Zvetkova I., Appanah R., Nesterova T.B., Webster Z., Peters A.H.F.M., Jenuwein T., Otte A.P., and Brockdorff N. 2003. Establishment of histone H3 methylation on the inactive X chro-

mosome requires transient recruitment of Eed-Enx1 polycomb group complexes. *Dev. Cell* **4:** 481–495.

Smith K.P., Byron M., Clemson C.M., and Lawrence J.B. 2004. Ubiquitinated proteins including uH2A on the human and mouse inactive X chromosome: Enrichment in gene rich bands. *Chromosoma* **113:** 324–335.

Tada M., Takahama Y., Abe K., Nakatsuji N., and Tada T. 2001. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol.* **11:** 1553–1558.

Tada T., Obata Y., Tada M., Goto Y., Nakatsuji N., Tan S.S., Kono T., and Takagi N. 2000. Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth. *Development* **127:** 3101–3105.

Takagi N. and Sasaki M. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256:** 640–642.

Turner J.M., Mahadevaiah S.K., Fernandez-Capetillo O., Nussenzweig A., Xu X., Deng C.X., and Burgoyne P.S. 2005. Silencing of unsynapsed meiotic chromosomes in the mouse. *Nat. Genet.* **37:** 41–47.

Wakefield M.J., Keohane A.M., Turner B.M., and Marshall Graves J.A. 1997. Histone underacetylation is an ancient component of mammalian X chromosome inactivation. *Proc. Natl. Acad. Sci.* **94:** 9665–9668.

White W.M., Willard H.F., Van Dyke D.L., and Wolff D.J. 1998. The spreading of X inactivation into autosomal material of an X; autosome translocation: Evidence for a difference between autosomal and X chromosomal DNA. *Am. J. Hum. Genet.* **63:** 20–28.

Wutz A. and Jaenisch R. 2000. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol. Cell* **5:** 695–705.

Wutz A., Rasmussen T.P., and Jaenisch R. 2002 Chromosomal silencing and localization are mediated by different domains of *Xist* RNA. *Nat. Genet.* **30:** 167–174.

# 第 18 章

# 哺乳动物的 DNA 甲基化


En Li[1] and Adrian Bird[2]

[1]*Novartis Institutes for BioMedical Research, Inc., Cambridge, Massachusetts 02139*

[2]*The Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh, EH9 3JR, United Kingdom*


## 本章目录


1. 细胞记忆机制
   - 1.1 假说
   - 1.2 甲基化模式可继承的证据
   - 1.3 哺乳动物的 DNA 维持甲基转移酶
2. DNA 甲基化模式的起源
   - 2.1 早期胚胎中 DNA 从头甲基化
   - 2.2 从头甲基转移酶的发现
   - 2.3 CpG 岛和 DNA 甲基化模式
   - 2.4 发育过程中 DNA 甲基化模式的动态变化
   - 2.5 受精卵中父本基因组的主动去甲基化
   - 2.6 什么使 CpG 岛不受 DNA 甲基化影响?
   - 2.7 染色质结构触发 DNA 甲基化?
   - 2.8 SWI/SNF-like 染色质重塑蛋白的作用
3. DNA 甲基化对基因表达的调控
   - 3.1 早期证据
   - 3.2 转录因子结合的干扰
   - 3.3 甲基化 CpG 结合蛋白的吸引
   - 3.4 MeCP2 和 Rett 综合征
   - 3.5 MBD2 调控依赖于甲基化的转录抑制
4. DNA 甲基化、突变和染色体的稳定性
   - 4.1 DNA 甲基化和突变
   - 4.2 DNA 甲基化和染色体不稳定性
5. 展望
   - 5.1 诱导表观遗传改变的外因
   - 5.2 表观遗传不稳定和复杂疾病
   - 5.3 可逆表观遗传状态的调节

参考文献

## 概要

脊椎动物DNA对5′CpG3′二核苷酸序列的胞嘧啶碱基发生甲基化共价修饰。CpG是DNA中被磷酸相连的胞嘧啶和鸟嘌呤的缩写。哺乳动物在胚胎发育时期建立DNA甲基化模式，在细胞分裂时通过复制维持这种甲基化状态。DNA甲基化模式的可继承性使基因组的表观遗传标记在大量的细胞分裂过程中得以稳定，从而建立了一种细胞记忆的形式。

分子和遗传学研究都表明DNA胞嘧啶甲基化与基因沉默有关，并且在诸如X染色体失活、基因印记等事件中起重要作用。甲基位于DNA双螺旋的大沟内，许多DNA结合蛋白在这里与DNA结合，甲基就是通过吸引或排斥各种DNA结合蛋白而发挥其作用。能够结合在甲基化的CpG二核苷酸的蛋白家族被称作甲基化CpG结合蛋白，它们能够募集抑制因子到发生了甲基化的启动子区域从而引起转录水平沉默。某些转录因子只与CpG未甲基化的DNA序列结合，这时CpG若发生甲基化就可以阻止这些蛋白结合并影响转录。

对编码DNA甲基转移酶和甲基化CpG结合蛋白的基因敲除（genetic ablation）证明了DNA甲基化在哺乳动物发育中起到多种作用。基因组的正常甲基化模式的建立对胚胎发育是至关重要的。维持印记基因的父本和母本拷贝的差量表达以及失活的那条X染色体上基因的持续沉默都需要DNA甲基化。另外DNA甲基化对于维持前病毒（proviral）基因组和内源逆转座子的转录抑制也是必不可少的。我们已经有了在发育过程中组织特异性基因表达模式的建立和维持离不开DNA甲基化的例子，同时DNA甲基化的缺失导致染色体数目不稳定，使染色体丢失的频率增加。

DNA甲基化的临床意义首先在于与癌症的联系。通过基因手段或DNA甲基转移酶抑制剂治疗降低DNA甲基化水平使某些小鼠模型的肿瘤得到了抑制，相反，一些其他类型的肿瘤由于DNA甲基化水平的降低而增强了。在人类其他疾病中也发现了编码关键的DNA甲基化装置组分的基因突变。DNA甲基转移酶3b（DNA methyl-transferase 3b，Dnmt3b）的突变导致免疫缺陷，甲基化CpG结合蛋白2（methyl-CpG-binding protein 2，MeCP2）的突变引发严重的神经疾病——Rett综合征。由此可见DNA甲基化系统的完整对于哺乳动物的健康是极为重要的。

虽然DNA甲基化模式可以在细胞间传递，但它不是永久的。实际上，个体的一生中都发生着DNA甲基化模式的改变。一些变化可能是环境改变的生理反应，另外一些变化可能与恶性转化或细胞老化等病理进程有关。至今我们还不是很清楚导致DNA甲基化改变的内外因素。对人类疾病的DNA甲基化研究是医学上的重要前沿，将有助于我们理解表观遗传对人类生命的影响。

## 1. 细胞记忆机制

### 1.1　假说

在哺乳细胞中胞嘧啶甲基化主要发生在 CpG 二核苷酸上（图 18-1）。动物体内 DNA 甲基化象征着一种细胞记忆机制的观点由两个实验室分别独立提出（Holliday and Pugh 1975；Riggs 1975）。这两个实验小组在意识到 CpG 二核苷酸是自我互补后，同时推论 CpG 的甲基化和非甲基化模式能够在细胞分裂时得到复制。在 DNA 刚刚复制完成时，来自亲本的 DNA 链继续维持它被修饰的胞嘧啶的模式，而新合成的那条链是没有被修饰的。为了使亲本的模式传递给子代，他们假定有种“维持甲基转移酶”能够专一地利用甲基化的亲本 CpG 为模板使子链的 CpG 甲基化，而非甲基化的 CpG 不能成为维持甲基转移酶的模板（图 18-2）。这种简单的机制产生的结果是 DNA 甲基化模式可以像 DNA 序列本身一样被半保留复制。

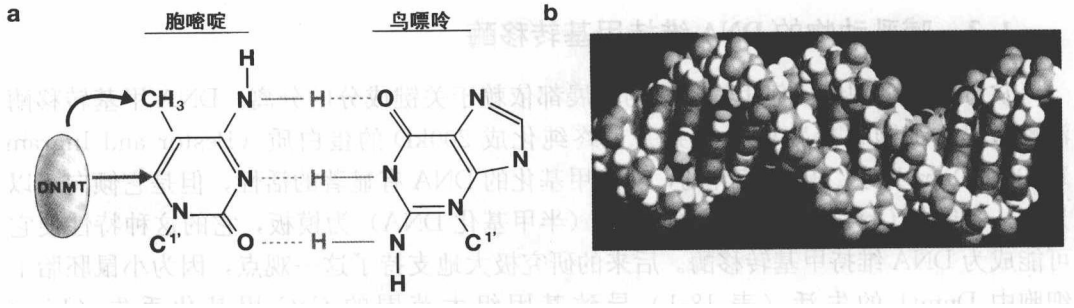



图 18-1　DNA 中的胞嘧啶甲基化

(a) 一个甲基（红色）连在胞嘧啶的嘧啶环（黑色箭头）5′位置上不会影响 GC 配对（蓝色）。在甲基转移过程中，DNA 甲基转移酶共价结合在 C6 位上（绿色箭头）。(b) B 型 DNA 模型中两个互补的 CpG 序列上的胞嘧啶被甲基化，成对的甲基位于双链的大沟内。

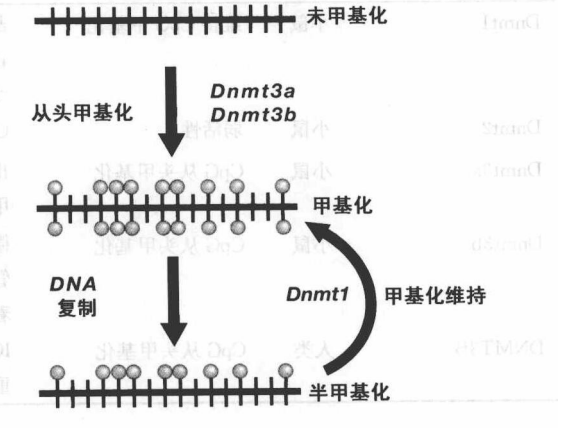



图 18-2　DNA 的从头甲基化和维持甲基化

图中所示水平线代表 1 段基因组 DNA，垂直的线条代表互补的 CpG。上图的非甲基化 DNA 在 Dnmt3a 和 Dnmt3b 的作用下，某些位置的 CpG 发生对称的甲基化，即为从头甲基化。由于 DNA 复制的半保留特征，后代 DNA 的一条链就与甲基化的母链配对（另一条复制后代链没有显示），维持 DNA 甲基转移酶 Dnmt1 只能以一条甲基化的链为模板使互补链甲基化，从而使甲基化的对称保留下去。

## 1.2 甲基化模式可继承的证据

最早人们是利用对 DNA 甲基化敏感的限制性内切核酸酶建立了单一基因组位点的 DNA 甲基化模式。如果特异的碱基发生了甲基化，那么这种酶就不能切割它们识别的序列。人们利用这种能切割胞嘧啶未发生甲基化但不能切割胞嘧啶甲基化的 CpG 位点的酶，绘出了非洲爪蟾核糖体 RNA 基因甲基化和非甲基化的模式图（Bird and Southern 1978）。他们发现，在一条特定的 DNA 链上，大部分的 CpG 是甲基化的，但是非甲基化的位点会随机出现，重要的是，这些随机出现的位点总是对称的，也就是说，两条链的对应 CpG 不是全部甲基化就是全部非甲基化（Bird 1978）。这个发现与之前维持甲基化模型的推测是十分吻合的（Holliday and Pugh 1975；Riggs 1975）。通过将人工合成的甲基化 DNA 转染进细胞为 DNA 甲基化模式的可遗传提供了更直接的检测途径（Wigler 1981），未甲基化的质粒即使经过很多代通常也不会发生甲基化，而 CCGG 位点被甲基化酶 M. HpaⅡ甲基化的质粒，经过很多代后仍保留它的甲基化状态，尽管培养细胞中该过程的保真度低于 100%。

## 1.3 哺乳动物的 DNA 维持甲基转移酶

从分子水平理解任何生物过程的进展都依赖于关键成分的分离。DNA 甲基转移酶活性最早在细胞粗提物中检测到并最终纯化成 200kD 的蛋白质（Bestor and Ingram 1983）。Dnmt1 对 CpG 是特异的，对非甲基化的 DNA 有显著的活性，但是它倾向于以只有一条链的 CpG 发生甲基化的 DNA（半甲基化 DNA）为模板，它的这种特性使它可能成为 DNA 维持甲基转移酶。后来的研究极大地支持了这一观点，因为小鼠胚胎干细胞中 Dnmt1 的失活（表 18-1）导致基因组大范围的 CpG 甲基化丢失（Li et al. 1992）。正如 Riggs、Holliday 和 Pugh 推测的那样，这些明显的证据都与 Dnmt1 通过使 CpG 半甲基化位点完全甲基化来维持 DNA 甲基化的观点相吻合（图 18-2）。

**表 18-1 DNA 甲基转移酶的功能**

| DNA 甲基转移酶 | 物种 | 主要活性 | 失活后的重要表型 |
|---|---|---|---|
| Dnmt1 | 小鼠 | 维持 CpG 甲基化 | 基因组大规模的 DNA 甲基化丢失，E9.5 胚胎致死，印记基因表达异常，异位 X 染色体的失活，沉默反向转座子的活化 |
| Dnmt2 | 小鼠 | 弱活性 | CpG 甲基化无改变，无显著发育表型 |
| Dnmt3a | 小鼠 | CpG 从头甲基化 | 出生后 4～8 周死亡，雄性不育，雌性和雄性无法建立甲基化印记 |
| Dnmt3b | 小鼠 | CpG 从头甲基化 | 微卫星 DNA 的去甲基化，E14.5 胚胎致死并且伴随血管和肝脏缺陷（Dnmt3a 和 Dnmt3b 同时缺失的胚胎在着床后无法启动从头甲基化而死于 E9.5） |
| DNMT3B | 人类 | CpG 从头甲基化 | ICF 综合征：免疫缺陷，着丝粒不稳定以及面部畸形；重复序列和近着丝粒异染色质的去甲基化 |

## 2. DNA 甲基化模式的起源

### 2.1　早期胚胎中 DNA 从头甲基化

我们需要了解的不仅包括 DNA 甲基化模式是怎样维持的，还包括它们最初是怎样产生的。早期的细胞转染实验表明，进入培养的体细胞的非甲基化 DNA 在数代后也通常会保留它的非甲基化状态。另一方面，导入小鼠着床前胚胎的逆转录病毒、前病毒和其他转基因在动物细胞内会稳定地被甲基化（Jahner et al. 1982），这暗示了 DNA 的从头甲基化过程被限制在胚胎发生的全能期，这一观点随后在小鼠胚胎瘤细胞和胚胎干细胞中得到验证。与体细胞不同，当向这些细胞中导入逆转录病毒时，逆转录病毒是没有传染性的，这也支持了病毒的传染周期（Stewart et al. 1982），另外，这些前病毒在 CpG 处发生了甲基化，把它们甲基化的基因组克隆到细菌中而去除 DNA 甲基化后，这些病毒仍然能进行基因表达。这些结果表明 DNA 甲基化可以在体内沉默病毒基因组的表达，而且胚胎细胞的确具有使 DNA 从头甲基化的能力。由于最初只知道 Dnmt1 一种 DNA 甲基转移酶，因此它被认为具有在发育期使 DNA 从头甲基化的能力，然而，Dnmt1 基因的去除并没有影响胚胎干细胞中前病毒 DNA 从头甲基化（Lei et al. 1996），这说明起作用的应该是其他的 DNA 甲基转移酶。

### 2.2　从头甲基转移酶的发现

原核细胞所有的胞嘧啶 DNA 甲基转移酶都具有特征相同的结构域（Posfai et al. 1989），而 DNA 维持甲基转移酶 Dnmt1 也具有这个特点。通过表达序列标签（expressed sequence tag，EST）数据库的搜索找到了三个可能编码其余 DNA 甲基转移酶的转录物（图 18-3）。其中一个候选蛋白 Dnmt2 在体外的 DNA 甲基转移酶活性极低，它的缺失对 DNA 甲基化水平没有影响。另外两个候选蛋白是 Dnmt3a 和 Dnmt3b，编码有催化活性的相关蛋白，但是与 Dnmt1 不同的是，它们在体外对半甲基化 DNA 无偏好（Okano et al. 1998）。缺失 Dnmt3a 和 Dnmt3b 的小鼠胚胎干细胞和胚胎不能使前病毒基因组和重复元件从头甲基化（Okano et al. 1999），这表明这两个蛋白是从头甲基转移酶（表 18-1）。此外，缺少 Dnmt3a 或相关调控因子 Dnmt3L 的雌性和雄性生殖细胞不能在印记基因上建立不同的甲基化模式（表 18-1）（Hata et al. 2002；Kaneda et al. 2004），Dnmt3a 和 Dnmt3b 同时失活导致早期胚胎死亡，单独一个基因失活导致严重缺陷，出生后死亡（Dnmt3a）或胚胎死亡（Dnmt3b）。*Dnmt3b* 基因突变引起一种以免疫缺陷、着丝粒不稳定和面部异常为特征的罕见疾病——ICF 综合征（The Immunodeficiency，Centromeric region instability，Facial anomalies syndrome）（Ehrlich 2003），这为从头甲基转移酶在人类中的作用提供了证据。对从 ICF 患者体内分离的淋巴样干细胞 DNA 分析显示，基因组甲基化水平下降，尤其是与染色体近着丝粒相连的 DNA 重复序列。

### 2.3　CpG 岛和 DNA 甲基化模式

哺乳动物成体组织的 DNA70%的 CpG 位点是甲基化的（Ehrlich 1982）。定位研究

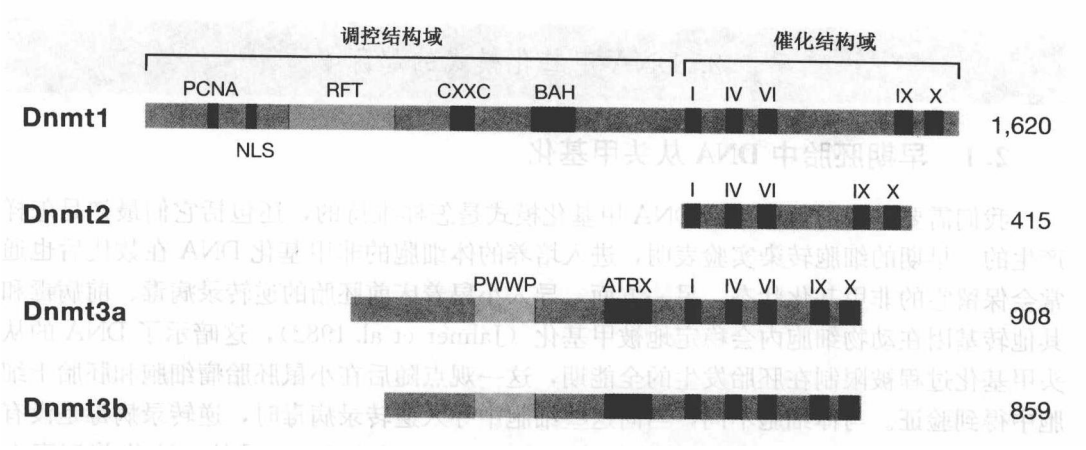



图 18-3 哺乳动物 DNA 甲基转移酶

Dnmt1、Dnmt2 和 Dnmt3 家族成员的催化结构域是保守的（标注Ⅰ、Ⅳ、Ⅵ、Ⅸ和Ⅹ的区域是在所有胞嘧啶甲基转移酶中最保守的），但它们的 N 端调控结构域的相似度很低。PCNA，PCNA 结合结构域；NLS，nuclear localization signal，核定位信号；RFT，replication foci-targeting domain；CXXC，半胱氨酸富集区，能够结合包含 CpG 双核苷酸的 DNA 序列；BAH，bromo-adjacent homology 结构域，与蛋白之间相互反应有关；PWWP，包含高度保守的“脯氨酸-色氨酸-色氨酸-脯氨酸”的结构域，与异染色质的联结有关；ATRX，与 ATRX 相关的半胱氨酸富集区，包含一个 C2-C2 锌指结构和非典型的 PHD 区，与蛋白间作用有关。

（见 Box 1）表明，卫星 DNA、包括转座子和其惰性残留（inert relics）的重复元件、非重复基因间 DNA 和外显子都是高甲基化的序列。在这些序列中，似乎对甲基化某种序列没有特定的偏好。卫星 DNA 和转座元件或外显子的 CpG 甲基化程度都是一样的，因此，大部分序列是根据能反映它们碱基组成的 CpG 二核苷酸的频率被甲基化的。哺乳动物基因组这种整体甲基化的例外是 CpG 岛，CpG 岛非同寻常地易被 DNA 甲基化敏感的限制性内切酶切割（Cooper et al. 1983），它是作为脊椎动物 DNA 片段被发现的。对所谓的“HpaⅡ微小片段”的克隆显示，它们是从长约 1kb 的富含 GC 的序列中分离出来的，这些序列在生殖细胞、早期胚胎和通常所有成体组织中都是未甲基化的（Bird et al. 1985），因此，CpG 岛是存在于大多数哺乳动物基因组中的“整体（global）”CpG 甲基化的例外（图 18-4）。早期个别基因启动子的定位认为 GC 富集区域靠近启动子（McKeon et al. 1982），现在基本上把大多数 CpG 岛标记于启动子或 5′区域，大约 60%的人类基因的启动子是含有 CpG 岛的。

### Box 1 DNA 甲基化定位

为了解 DNA 甲基化的功能，首要应该做的是找出它们在基因组中发生的位置。目前已有几种方法进行此项工作，各有利弊。

- 由于甲基化主要发生在 CpG 序列上，因此能识别含 CpG 的 DNA 序列的限制性内切酶被广泛用于甲基化定位（Bird and Southern 1978）。应用该方法能够分析基因组中的大部分区域，但只局限于限制性内切酶识别位点的 CpG。
- 单链 DNA 的亚硫酸氢盐修饰是一种能检测某一区域所有胞嘧啶状态的可靠方法（From-

mer et al. 1992)。亚硫酸氢盐能使未修饰的胞嘧啶脱氨基，而 5-甲基化的胞密啶不会脱氨基，因此经过亚硫酸氢盐处理后保留下来的胞嘧啶被认为是甲基化的。由于甲基化胞嘧啶的高分辨率和阳性识别率，该方法被用来分析 DNA 的甲基化模式，但是目前还不能用于大区域的彻底分析。

- 建立在亚硫酸氢盐处理的基础上又有几种通过 PCR 的技术被用来分析感兴趣的区域(Herman et al. 1996)。这些方法简便易行，但由于集中于某一区域的部分 CpG 位点，因此需通过测序得到详细的信息。
- 最近芯片被用于绘制 DNA 甲基化图谱。例如，在布满基因组 DNA 序列的芯片上，5-甲基化胞嘧啶特异的核酸酶无法消化的 DNA 能被标记出来，这就能从总体上观测到特异区域的甲基化水平（Martienssen et al. 2005)，应用 5-甲基化胞嘧啶特异的抗体还能对这些探针进行免疫沉淀，从而对 DNA 甲基化水平有了整体检测（Weber et al. 2005)。

## 2.4　发育过程中 DNA 甲基化模式的动态变化

在各种不同的成体细胞中 DNA 甲基化模式表现出整体的稳定性，但是在某些特异 DNA 序列上也会发生显著的变化，例如，雌性哺乳动物在 X 染色体失活的胚胎期，一条 X 染色体上的 CpG 岛会大量地发生从头甲基化（Wolf et al. 1984)，这一过程对于防止失活 X 染色体上的基因沉默是十分重要的（图 18-4)，因为 X 染色体相关基因在 DNA 甲基化缺陷的小鼠或细胞内重新转录活化的频率极高。分化过程中某些基因的转录活化也被发现与 DNA 甲基化的程序化丢失有关，例如，在 T 细胞分化过程中，白细胞介素 2 (interleukin-2) 基因表达时在启动子区丢失了 CpG 甲基化（Bruniquel and Schwartz 2003)，这个去甲基化事件是基因活化的重要前提，因此是 T 细胞分化进程的关键部分。

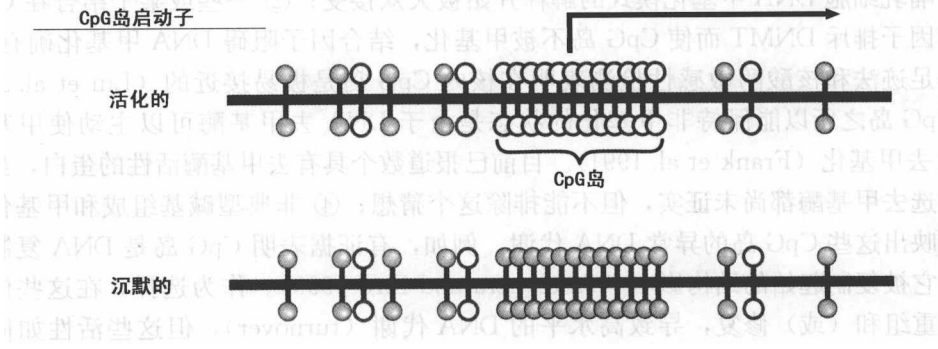



图 18-4　CpG 岛

CpG 岛出现在大部分人类基因的启动子区的，其富含 CpG，且通常是非甲基化的序列。CpG 岛的甲基化可以保证这些基因的长期沉默，例如，失活的 X 染色体上的基因和某些印记基因就是通过这条途径发生沉默，另外癌细胞中 CpG 岛的甲基化也使某些基因发生异常沉默。

## 2.5　受精卵中父本基因组的主动去甲基化

在受精卵中除了发生 DNA 甲基化的这些局部变化，总体 DNA 甲基化水平也发生

巨大的改变。最初利用5-甲基胞嘧啶抗体对DNA甲基化水平的分析表明，在早期胚胎分裂时期一整套染色体的DNA甲基化是高度缺失的（Rougier et al. 1998）。这些差别通过对5-甲基胞嘧啶免疫染色检测母本与父本核融合前的合子而发现（Mayer et al. 2000）。刚开始时，母本和父本的原核显示出相同的着色，这表示卵子和精子的基因组的胞嘧啶甲基化水平大致相同，然而在受精后几个小时，来自父本基因组的5-甲基胞嘧啶水平发生显著的下降，亚硫酸氢盐测序也证实其中一个基因组是去甲基化的。虽然DNA甲基化丢失的机制还不清楚，但是它必然包含修饰的“主动”切除，因为在该时期并没有发生DNA复制。有趣的是，在早期胚胎发生过程中母本基因组也丢失了DNA甲基化，但由于早期卵裂时不存在维持DNA甲基化因此这个过程是“被动”的，总之在这一过程中超过半数的DNA甲基化丢失了，在着床期才发生依赖于Dnmt3a和Dnmt3b的从头甲基化。

除此之外，还有其他DNA主动去甲基化的报道。例如，在分化的辅助T细胞中，白介素2基因启动子在没有DNA复制的情况下迅速发生去甲基化（如前述和Bruniquel and Schwartz 2003）。与此相当的效应在蛙卵母系细胞系统中人为地得以再现。该系统中沉默的甲基化的哺乳动物基因*Oct-4*能够经过依赖于基因的优先去甲基化的过程而重新转录激活（Simonsson and Gurdon 2004）。

## 2.6 什么使CpG岛不受DNA甲基化影响？

DNA甲基化模式的研究一直集中于哺乳细胞CpG岛如何不受总体DNA甲基化的影响。最简单的可能的解释是：① 已有的DNA从头甲基转移酶根本无法使CpG岛甲基化，但这好像是不可能的，因为同一细胞中失活的那条X染色体上的CpG岛被大量甲基化，而有活性的X染色体上的CpG岛保持非甲基化。另外，在癌细胞和癌细胞系中，许多正常情况下非甲基化的CpG岛发生了甲基化，因此其他几种不存在排他性的关于哺乳细胞DNA甲基化模式的解释开始被大众接受；② 一些或某个结合在CpG岛上的因子排斥DNMT而使CpG岛不被甲基化，结合因子阻碍DNA甲基化确有证据，但是足迹法和核酸酶敏感性检测表明在核内CpG岛是极易接近的（Lin et al. 2000）；③ CpG岛之所以能保持非甲基化的状态是由于DNA去甲基酶可以主动使甲基化的CpG去甲基化（Frank et al. 1991）。目前已报道数个具有去甲基酶活性的蛋白，虽然这些候选去甲基酶都尚未证实，但不能排除这个猜想；④ 非典型碱基组成和甲基化的缺乏反映出这些CpG岛的异常DNA代谢。例如，有证据表明CpG岛是DNA复制的源头，它被复制起始的结构影响（Antequera and Bird 1999）。作为选择，在这些位点集中了重组和（或）修复，导致高水平的DNA代谢（turnover），但这些活性如何排斥DNMT还不清楚；⑤ DNA甲基化被排斥需要含CpG岛的启动子在早期胚胎转录来保证，在转基因检测中启动子的突变促进CpG岛甲基化（Brandeis et al. 1994；MacLeod et al. 1994），高组织特异性表达基因的含CpG岛的启动子通常在早期胚胎中是表达的，但是现在还缺少转录能排斥CpG甲基化的证据。

## 2.7 染色质结构触发DNA甲基化？

上述的假说认为CpG甲基化是基因组的默认状态，CpG岛是由于被排除在总体甲

基化外而产生的。最近出现了一些不同的观点：潜在的染色质修饰状态决定 DNA 甲基化模式，由此在更早期通过特异途径对组蛋白的修饰触发了 DNA 的甲基化或非甲基化，这一观点在链胞霉和拟南芥中得到了强有力的证据（见第 6、9 章）。在这些系统中，甲基化并没有局限在 CpG 位点，而是依赖于组蛋白 H3K9 甲基化（H3K9me）的出现（Tamaru and Selker 2001；Jackson et al. 2002）。与此相关的是，在植物中已有人认为 RNAi 对染色质修饰、基因沉默和 DNA 甲基化起作用（Aufsatz et al. 2002；第 8 章），但是在哺乳动物中，这种联系还没有很好的证据。哺乳动物中 CpG 甲基化是主要的，两种特异作用于 H3K9 位点的组蛋白赖氨酸甲基转移酶（HKMT）的缺失，能够降低异染色质重复序列内的 CpG 甲基化（Lehnertz et al. 2003）。另外，敲除一种特异作用于 H3K37 位点的 HKMT——polycomb 蛋白 EZH2，能够引起 EZH2 的目标启动子 CpG 甲基化的丢失（Vire et al. 2006）。哺乳动物中反义转录触发 DNA 甲基化也有强力的证据。Higgs 及其合作者（Tufarelli et al. 2003）得到的结果显示，在分化的小鼠胚胎细胞中通过 α 珠蛋白基因驱动的反义转录产物能够保证 CpG 岛甲基化，虽然这一效应的机制还不清楚，但推测 RNAi 在其中是起作用的，因为在植物中它确实参与了从头甲基化（Zilberman et al. 2003）。RNAi 触发从头甲基化在培养的哺乳细胞中也有报道（Kawasaki and Taira 2004），但目前的数据暗示哺乳动物中这一现象不如植物和霉菌中清晰。

### 2.8　SWI/SNF-like 染色质重塑蛋白的作用

在适当甲基化过程中需要染色质附属因子（accessory factor）的证据源于对植物的研究。植物的拟南芥中 SNF2 样蛋白被证明对基因组的完全甲基化是必需的（Jeddeloh et al. 1999）。动物中也存在相同的依赖性，因为人 *ATRX* 基因（alpha thalassemia/mental retardation syndrome X-linked）（Gibbons et al. 2000）和小鼠 *Lsh2* 基因（helicase lymphoid specific）（Dennis et al. 2001）的突变对总体 DNA 甲基化模式有巨大影响，而这两个基因都是编码与染色质重塑蛋白 SNF2 相关的蛋白。尤其是 LSH2 蛋白的缺失与拟南芥 DDM1 突变导致的表型吻合，二者的突变体都丢失了高重复序列的甲基化而基因组其他位置的甲基化没有改变。基因组有效的甲基化也许需要这些染色质重塑蛋白引发的染色质结构混乱，从而使 DNMT 能够接近 DNA，DNMT 和染色质因子之间的协作使它们进入特异的染色质区域，这对异染色质和难以靠近的区域可能尤其重要。

## 3. DNA 甲基化对基因表达的调控

### 3.1　早期证据

人们已经用几种方法检测了 DNA 甲基化对基因表达调控的影响。利用 M. HpaⅡ人为甲基化腺病毒报告基因的部分 CpG 位点，当被注射到蛙卵母细胞核时，可以抑制基因表达（Vardimon et al. 1982）。与此类似的是，当转染到培养的哺乳动物细胞中时，腺嘌呤磷酸核糖基转移酶基因也由于 CpG 甲基化而沉默（Stein et al. 1982）。由于 5-氮杂胞苷（5-azacytidine）可以抑制活体细胞 DNA 甲基化（Jones and Taylor 1980），使

DNA 甲基化对自然状态的基因组 DNA 影响的研究成为可能，这种核苷类似物进入 DNA 代替了胞嘧啶，并与 DNA 甲基转移酶形成共价复合物，使它不能继续甲基化 DNA。人们已经发现几种基因的沉默，包括病毒基因组（Harbers et al. 1981）和失活 X 染色体上的基因（Wolf et al. 1984）都与它们的甲基化相关，5-azacytidine 能够使它们恢复表达（Mohandas et al. 1981），这一结果表明 DNA 甲基化可能对它们的表达起决定性的作用，通过检测 5-azacytidine 处理过的经纯化的细胞提取物，人们证实产生这一效应的原因是 DNA 甲基化的变化而不是药物的其他影响。用 5-azacytidine 处理过的细胞 DNA 转染其他细胞时，失活 X 染色体相关的次黄嘌呤磷酸核糖基转移酶基因得到表达，而对照的没有用 5-azacytidine 处理过的细胞 DNA 不能使该基因表达（Venolia et al. 1982）。

### 3.2 转录因子结合的干扰

DNA 甲基化是如何干扰基因表达的？一种最明显的可能是 DNA 大沟中甲基的出现（图 18-1）阻碍了特定基因转录活化的转录因子的结合。许多转录因子识别包含 CpG 的 GC 富集序列，当 CpG 被甲基化时，其中一些转录因子就不能结合 DNA（Watt and Molloy 1988），小鼠中 CTCF（CCCTC 结合因子）蛋白对 *H19/ Igf2* 印记作用的研究证明了这一机制确实参与基因表达调控（Bell and Felsenfeld 2000）。CTCF 结合在转录区域边界（Bell et al. 1999），它能使启动子不受远处增强子的影响，由于 CTCF 结合在 *Igf2* 启动子和下游增强子之间，因此来自母本的 *Ifg2* 基因拷贝是沉默的，然而在父本的基因拷贝中 CTCF 结合的 CpG 位点是甲基化的，因此 CTCF 不能结合从而使下游增强子激活 *Igf2* 的表达。虽然有证据显示 *H19/Ifg2* 印记效应的产生也包括其他原因，但是 CTCF 代表了 DNA 甲基化能够调控转录的显著例子（详见第 19 章）。

### 3.3 甲基化 CpG 结合蛋白的吸引

抑制的第二种模式与第一种正相反，因为它包含被甲基化 CpG 吸引而非排斥的蛋白（图 18-5 和表 18-2），在能够支持外来基因转录的哺乳细胞提取物中首先检测到了这种抑制模式。加入微量的 DNA 允许非甲基化的报告基因转录，而甲基化的报告基因会被抑制（Boyes and Bird 1991）；而增加外来 DNA 量则可引起甲基化和非甲基化的模板转录的水平相同，这着味着有限量的 DNA 甲基化特异的转录抑制剂被饱和。与这个解释相一致的是，在这些提取物中过量的甲基化非特异竞争 DNA 解除了对甲基化基因的抑制。导入哺乳动物细胞和蛙卵母细胞的甲基化 DNA 允许基因转录的实验为转录的间接抑制提供了证据（Buschhausen et al. 1987；Kass et al. 1997），沉默发生于几个小时后意味着沉默可能依赖于染色质的装配。为找到与甲基化基因特异结合的蛋白并产生可观察到抑制作用，人们利用随机甲基化 DNA 序列作为探针进行凝胶阻滞实验，在多种哺乳细胞系中找到了一种特异于甲基化 DNA 的 DNA-蛋白复合体（MeCP1）（Meehan et al. 1989），然而第一个被纯化克隆出的甲基化 CpG 结合蛋白是 MeCP2。利用数据库已经鉴定出了几个甲基化 CpG 结合结构域（methyl-CpG-binding domain，MBD）蛋白家族成员，分别是 MeCP2、MBD1、MBD2、MBD3 和 MBD4（表 18-2）（Bird and

Wolffe 1999)，这些蛋白的 DNA 结合结构域与 MeCP2 的 DNA 结合结构域相关，其中 MBD2 是 MeCP1 复合体的 DNA 结合部分。

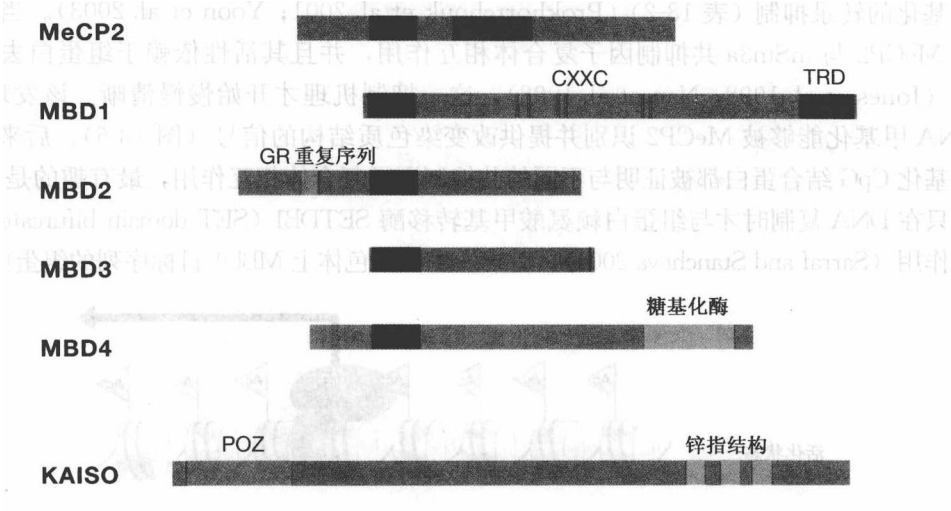



图 18-5　结合甲基化 CpG 的蛋白质

MBD 蛋白家族的 5 个成员按照它们的 MBD 结构域排列（紫色标注），其他区域包括转录抑制结构域（TRD）；CXXC 结构域（锌指结构域），某些锌指结构能与非甲基化的 CpG 结合；GR 重复区，未知功能；TG 错配的糖基化酶，与脱氨基的 5-甲基化胞嘧啶的修复有关。Kaiso 缺少 MBD 结构域，却能通过锌指（橘黄）结合甲基化 DNA，并且与其他转录抑制因子共有一个 POB/BTB 区域。

**表 18-2　甲基化 CpG 结合蛋白的功能**

| MBP | 主要活性 | 种属 | 失活后的重要表型 |
|---|---|---|---|
| MeCP2 | 结合相邻 AT 的甲基化 CpG；转录抑制子 | 小鼠 | 神经缺陷的延迟发作，包括惯性、后肢紧扣、非间歇性呼吸和异常步伐；出生后 10 周仍存活 |
| MECP2 | 结合相邻 AT 的甲基化 CpG；转录抑制子 | 人类 | 由于 Rett 综合征而出现异质性，显著的神经紊乱，主要是运用失能、呼吸和头小畸形 |
| Mbd1 | 通过 MBD 结合甲基化 CpG；其中一个剪切体还能通过 CxxC 区域结合 CpG | 小鼠 | 无显著表型，但有神经发育的微弱缺陷 |
| Mbd2 | 结合甲基化 CpG；转录抑制子 | 小鼠 | 能够繁育，但母性养育行为减弱；调节 T 辅助细胞分化的基因产生缺陷，因而导致对感染的反应改变；对肠癌发生有高抵抗力 |
| Mbd3 | NuRD 共抑制子复合物的中心组分；与甲基化 CpG 无强结合力 | 小鼠 | 早期胚胎致死 |
| Mbd4 | DNA 修复蛋白，与甲基化 CpG 结合，且在甲基化 CpG 位点发生 T:G 错配 | 小鼠 | 能够繁育；CpG 位点的突变率增加 3～4 倍；肠癌易感性增高，伴随着 Apc 基因内 C-T 转换；Mbd4 蛋白能将 5 甲基化胞嘧啶的突变减到最低 |
| Kaiso | 结合 mCGmCG 和 CTGCNA；转录抑制子 | 小鼠 | 无显著表型；对 Min 小鼠的肿瘤发生有微弱但重要的延缓 |

MBD1、MBD2 和 MeCP2 已经被证明参与依赖于甲基化的转录抑制过程（Bird and Wolffe 1999）。一个非相关蛋白 Kaiso 也在模式系统中结合于甲基化 DNA，引起依赖于甲基化的转录抑制（表 18-2）（Prokhortchouk et al. 2001；Yoon et al. 2003）。当人们发现 MeCP2 与 mSin3a 共抑制因子复合体相互作用，并且其活性依赖于组蛋白去乙酰化后（Jones et al. 1998；Nan et al. 1998），这一抑制机理才开始慢慢清晰。该发现表明，DNA 甲基化能够被 MeCP2 识别并提供改变染色质结构的信号（图 18-6）。后来这四个甲基化 CpG 结合蛋白都被证明与不同的共抑制因子复合体相互作用，最有趣的是 MBD1，它只在 DNA 复制时才与组蛋白赖氨酸甲基转移酶 SETDB1（SET domain bifurcated 1）相互作用（Sarraf and Stancheva 2004），后续发生在染色体上 MBD1 目标序列的组蛋白 H3K9

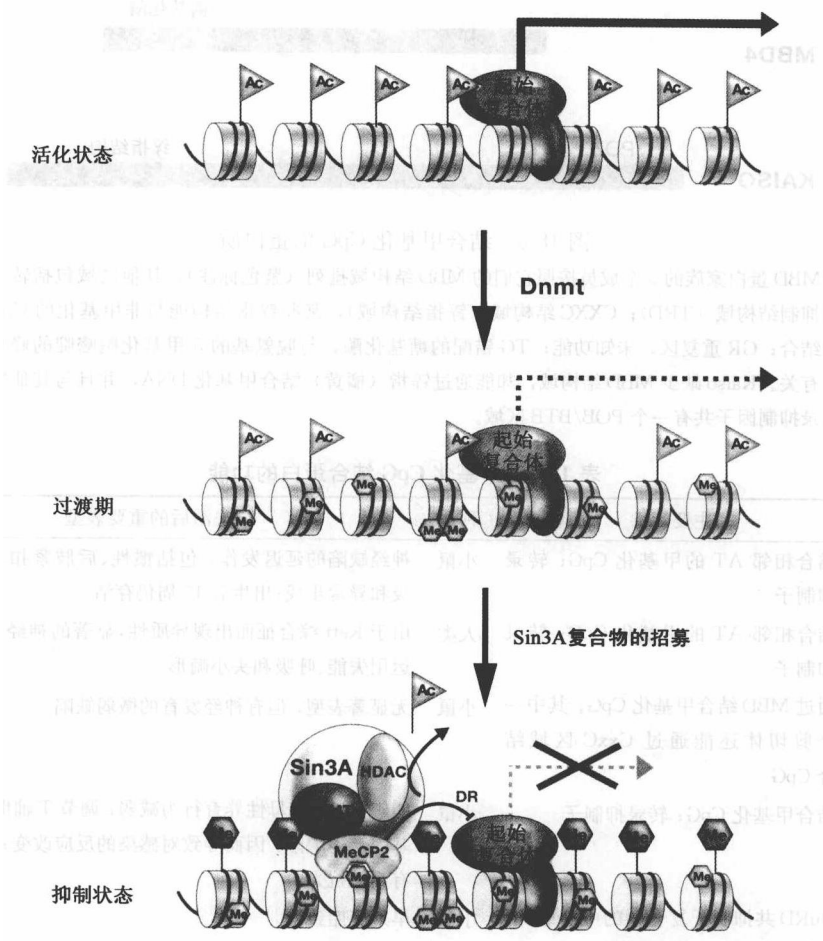



图 18-6 通过甲基化 CpG 结合蛋白募集辅抑制物

在 MeCP2 调节下，DNA 甲基化可能会使一个启动子由非甲基化的活化态变成甲基化的失活态。在这一过渡期中，转录的沉默和 DNA 甲基化都是迅速发生的。Sin3A 组蛋白去乙酰化酶（HDAC）和组蛋白赖氨酸甲基转移酶（HKMT）被认为是通过 MeCP2 募集到甲基化位点上的，另有证据表明 MeCP2 能通过与转录起始复合体 DR 相互作用而直接抑制转录。其他甲基化 CpG 结合蛋白也能够与各种辅抑制物包括 HKMT 和 HDAC 相互作用或者募集它们到甲基化位点。

甲基化和相关基因的稳定沉默，都依赖于这个染色质修饰反应的周期性招募。

要理解 MBD 蛋白生物学功能的首要条件是找到它们在基因组上的靶位点，人们推测它们识别的 DNA 序列特征有重叠，因此它们会互相竞争同一个甲基化位点。可令人惊讶的是，在原代培养的人细胞系中，MBD 结合位点在基因组中是基本不重叠的，这暗示着每个甲基化 CpG 结合蛋白是被独立定位的（Klose et al. 2005），最新证据显示 MBD 的结合有 DNA 序列特异性（表 18-2），这两者是吻合的。例如，MeCP2 对周围有一段 AT 富集序列的甲基化 CpG 位点有极大的偏好（Klose et al. 2005）；另外，MBD1 包含其他的 DNA 结合结构域，这个结构域是特异于非甲基化 CpG 的，而 Kaiso 识别一对相邻的甲基化 CpG 基序（Prokhortchouk et al. 2001），目前似乎只有 MBD2 仅对甲基化 CpG 有亲和性。

### 3.4　MeCP2 和 Rett 综合征

多种具有抑制因子性质的甲基化 CpG 结合蛋白的存在使人们推测，这些蛋白可能是甲基化信号的调节者。当人们发现一种严重的神经疾病-Rett 综合征（RETT）与 MECP2 基因的突变有关时，这一推测得到了极好的阐释。具有 X 染色体连锁的 MECP2 基因新型突变的杂合子的女性易感 RETT（表 18-2）（Amir et al. 1999）。由于 X 染色体的失活是随机的，因此 RETT 患者是混杂地表达突变型或野生型基因。这类女孩在 6～18 个月期间表现正常，但是之后她们会表现出运动技能受损、重复手部活动、呼吸异常、头小畸形等症状（表 18-2）；而具有同样突变的半合子男性不能存活。*Mecp2* 失活的小鼠在出生后几个星期内正常，但在第六个星期表现出一系列神经病症状，大约十个星期死亡。这个延期发作表型的几个特征类似人 Rett 综合征（Amir et al. 1999）。由于仅在小鼠脑中敲除 *Mecp2* 基因和在小鼠全身敲除引发的症状相同（表 18-2）（Chen et al. 2001；Guy et al. 2001），因此，虽然 MeCP2 在小鼠体内广泛地表达，但是 *Mecp2* 缺失小鼠的表型似乎是由于其在脑中的不表达的原因所致生物化学和免疫细胞化学研究也得到了同样的结果：MeCP2 在脑中表达水平最高，尤其是神经细胞，而且仅在神经细胞中表达 MeCP2 就能抑制小鼠表型的出现（Luikenhuis et al. 2004）。

既然 MeCP2 的功能是作为一个转录抑制因子，因此关于 RETT 产生了一个很有吸引力的假设：在脑中一些需要 MeCP2 抑制表达的基因，在缺乏 MeCP2 时得到表达因而导致异常的神经功能。哺乳动物中第一个被鉴定出来的 MeCP2 靶基因编码来源于脑的如神经因子 Bdnf（Chen et al. 2003；Martinowich et al. 2003）。Bdnf 属于需经神经活动诱导才合成的蛋白，它对于大脑活动中暂时的刺激向长期的变化这一转化过程十分重要，因此它的错误调控包含在 RETT 的病理进程中。有报道小鼠中其他几个基因的表达也发生变化，但产生的影响很小（小于 3 倍），然而这些变化中的一个或几个也可能与 RETT 表型有关。

### 3.5　MBD2 调控依赖于甲基化的转录抑制

MBD2 是 MeCP1 的 DNA 结合组分（表 18-2），MeCP1 最初被认为是细胞提取物中的转录抑制因子（Meehan et al. 1989；Boyes and Bird 1991），它是一个巨大的多蛋白复合体，

包含 NuRD（或 Mi-2）共制剂复合物和 MBD2（Wade et al. 1999）。NuRD 由组蛋白去乙酰化酶和一个大的染色质重塑蛋白（Mi-2）组成，通过 MBD2 和其他几个 DNA 结合蛋白，NuRD 被募集到 DNA 上，（有趣的是，与 MBD2 类似但不结合甲基化 DNA 的蛋白 MBD3，是 NuRD 复合体的一部分）。缺乏 MBD2 的细胞，即使具备其他甲基化 CpG 结合蛋白也不能有效地抑制甲基化的报告基因，这就说明 MBD2 是抑制系统的主要组分（Hendrich et al. 2001）。MBD2 缺失小鼠虽然母体行为异常，但是能够存活和繁育（Hendrich et al. 2001），而且细致的检测也发现组织特异基因表达的异常。白细胞介素 4 和干扰素 γ 基因的表达在辅助 T 细胞分化过程中也严重受损（Hutchins et al. 2002）。例如，相当多数的能够表达干扰素 γ 的 Th1 细胞也表达白细胞介素 4（表 18-2）。由于 MBD2 结合白细胞介素 4 基因，因此它的缺失可能削弱了 Th1 细胞中该基因的抑制。

## 4. DNA 甲基化、突变和染色体的稳定性

### 4.1 DNA 甲基化和突变

5-甲基化胞嘧啶易发生突变，这是反对将 DNA 甲基化作为细胞记忆的表观遗传系统的主要原因，胞嘧啶能自发脱氨基变成尿嘧啶，尿嘧啶会与鸟嘌呤配对（Lindahl 1974），尿嘧啶 DNA 糖苷酶能够识别这个潜在的突变，有效地移除不适当的碱基并启动将尿嘧啶替换成胞嘧啶的修复。然而如果甲基化的胞嘧啶自发脱氨基，就会形成胸腺嘧啶，虽然这也导致错配，但由于胸腺嘧啶是 DNA 中自然存在的碱基，因而不会被体内修复系统有效地修复，结果这个胞嘧啶到胸腺嘧啶的突变以 C-T 转换的形式通过 DNA 复制一直传给子代细胞。这个突变似乎是人类遗传疾病中最常见的单起因致病因素，因为几乎三分之一的点突变是 CpG 序列上的 C-T 转换（Cooper and Youssoufian 1988）。哺乳动物基因组的 CpG 含量仅占序列中 CpG 含量的四到五分之一，表明了在进化过程中 CpG 的不稳定性（Bird 1980）。然而 CpG 岛是唯一的例外，因为该 CpG 是非甲基化的并且是稳定的。

MBD4 是目前甲基化 CpG 结合蛋白中唯一具有酶活性的蛋白。它的 C 端结构域是胸腺嘧啶 DNA 糖苷酶，在体外能选择性地将 T-G 错配中的 T 移除（Hendrich et al. 1999）。人们希望这一活性能够修复 5-甲基化胞嘧啶脱氨基。与这个假设相一致的是，*MBD4* 缺失的小鼠在染色体报告基因上表现出更多的甲基化胞嘧啶位点突变（Millar et al. 2002），另外，*Mbd* 基因敲除的小鼠在腺瘤性息肉基因（adenomatous polyposis coli，APC）上有 C→T 突变，而且患肠癌的概率增加（表 18-2）。值得注意的是，尽管体内存在专门的修复系统，但胞嘧啶甲基化仍然是突变的热点。

### 4.2 DNA 甲基化和染色体不稳定性

虽然 DNA 甲基化是易突变的，但是有证据显示它的出现有利于染色体的稳定。小鼠由于 Dnmt1 的亚等位基因（hypomorphic）突变而 DNA 甲基化水平只有正常的 10%，这样的小鼠经常有三条 15 号染色体而发生浸润性的 T 细胞淋巴瘤（Gaudet et

al. 2003)。ICF 综合征的患者的 DNMT3B 突变或小鼠的 Dnmt3b 失活，都会导致各种染色体异常，包括染色体融合、断裂和非整倍性（Ehrlich 2003；Dodge et al. 2005），这些结果很有趣，因为癌症经常表现出与肿瘤发生和侵润相关的低甲基化水平，对此的一个可能的解释是 DNA 甲基化能够纠正染色体分离，而 DNA 甲基化的缺失使染色体融合的几率增加。DNA 甲基化能够抑制哺乳动物基因组中基因表达和逆转座子的重组，因而使染色体不会发生无益的重组。实际上，在胚胎发育和精子发生过程中，DNA 甲基化已经被证明在逆转座子的转录沉默中起关键作用（Walsh et al. 1998；Bourc'his and Bestor 2004)。

## 5. 展　望

虽然人们对哺乳动物中 DNA 甲基化生物学功能的了解越来越多，但目前这还远不够。例如，我们对于 CpG 甲基化的变异比例以及诱导甲基化模式改变的内因和外因并不像遗传突变那样清楚。越来越多的证据表明，DNA 甲基化和组蛋白修饰与许多复杂疾病的发病机理有关，因此基因组表观遗传状态的调节有可能成为治疗这些疾病的新途径，我们期待将来这些令人兴奋的研究领域出现新进展。

### 5.1　诱导表观遗传改变的外因

哺乳动物基因组的 DNA 甲基化模式在发育过程中是受高度调控的，而外在因素，如何影响 DNA 甲基化和基因表达还不清楚。最近已有研究表明外因是如何诱导长期生物效应的表观遗传改变。例如人们发现大鼠母亲的行为可以使后代的 DNA 甲基化产生稳定的变化，Weaver 和合作者报道收到母亲不同关怀的幼鼠在糖皮质激素受体（glucocorticoid receptor，GR）基因的启动子区域的甲基化是不同的，这种不同反过来与 GR 的表达正相关，而且这种不同抑制持续到成年（Weaver et al. 2004)。另外也有报道指出成年母鼠的饮食也会改变后代的 DNA 甲基化。*Agouti* 是小鼠的显性性状基因，赋予小鼠皮毛呈褐色（亦称 agouti 色）的特征。可转座的逆转录病毒元件插入到基因中能够产生几个豚鼠 AVY（*agouti viable yellow*）等位基因，在具有这种等位基因的小鼠中，*agouti* 的表达被逆转录病毒元件的长末端重复序列（long terminal repeat，LTR）控制，这些小鼠的皮毛不是有黄色斑点就是与野生型毛色相同，LTR 启动子的甲基化状态决定这种颜色变化，因此皮毛颜色成为 DNA 甲基化的一个外在表征。当喂食怀孕母鼠含有甲基化原料（如叶酸、胆碱和甜菜碱）的饲料时，它们的后代皮毛颜色就趋向于褐色，这与 LTR 启动子甲基化升高一致（Cooney et al. 2002；Waterland and Jirtle 2003)。这些结果都暗示了外在因素能够诱导甲基化状态的稳定改变，这或许是外因带来长期生物效应的机制。对于哺乳动物来说，表观遗传机制在基因与环境的相互作用中占了多大比重以及环境因素如何转化成表观状态，是未来需要解决的问题。

### 5.2　表观遗传不稳定和复杂疾病

许多复杂疾病，如Ⅱ型糖尿病、精神分裂症、自身免疫疾病和癌症等，虽然表现出

可遗传的特征，但是与孟德尔遗传规律并不十分相符。动态表观遗传机制为这些复杂疾病的某些特征提供了不同的解释，包括晚期发作、性别效应、父母本效应、同卵双生子的差异和症状的波动等（Petronis 2001）。虽然已有证据表明癌症与异常 DNA 甲基化和组蛋白修饰相关，但表观遗传在其他复杂疾病的病因中的作用还很不清楚，这些复杂疾病通常很难被检测到遗传突变，而正常人群和患者群之间基因组范围的 DNA 甲基化模式的比较，可以成为研究这些复杂疾病的表观遗传基础的切入点。

## 5.3 可逆表观遗传状态的调节

大部分表观遗传修饰是可逆的，这就使对表观遗传状态的调节成为治疗癌症和其他疾病的一个潜在的新选择。一些改变 DNA 甲基化模式或抑制组蛋白去乙酰化酶的药物已经在临床试验中得到检验（Egger et al. 2004），其中一些表现出抗癌效应。实际上，去甲基化药物 5-azacytidine 最近获得了美国食品与药物管理局的批准，用于治疗脊髓发育不良综合征。这是一种多相疾病，患者的造血细胞形态异常。5-azacytidine 和其他核苷酸类似物能够与 DNA 结合，其临床上的应用由于毒性受到限制，这鼓舞人们寻找能够直接抑制 DNA 甲基转移酶或作用于其他表观遗传调控因子的药物。由于 DNA 甲基化仅是复杂的表观遗传调控网络的一个部分，或许可以利用 DNA 甲基转移酶或组蛋白去乙酰化酶抑制剂来最大化治疗效果和最小化毒性，并与其他抗癌手段结合起来（详见第 24 章）。

赵 磊 译
王大亮 校
孙方霖 审校

## 参考文献

Amir R.E., Van den Veyver I.B., Wan M., Tran C.Q., Francke U., and Zoghbi H.Y. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23:** 185–188.

Antequera F. and Bird A. 1999. CpG islands as genomic footprints of promoters that are associated with replication origins. *Curr. Biol.* **9:** R661–667.

Aufsatz W., Mette M.F., van der Winden J., Matzke A.J., and Matzke M. 2002. RNA-directed DNA methylation in *Arabidopsis*. *Proc. Natl. Acad. Sci.* **99:** 16499–16506.

Bell A.C. and Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405:** 482–485.

Bell A.C., West A.G., and Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98:** 387–396.

Bestor T.H. and Ingram V.M. 1983. Two DNA methyltransferases from murine erythroleukemia cells: Purification, sequence specificity, and mode of interaction with DNA. *Proc. Natl. Acad. Sci.* **80:** 5559–5563.

Bird A.P. 1978. Use of restriction enzymes to study eukaryotic DNA methylation. II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J. Mol. Biol.* **118:** 48–60.

———. 1980. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8:** 1499–1594.

Bird A.P. and Southern E.M. 1978. Use of restriction enzymes to study eukaryotic DNA methylation. I. The methylation pattern in ribosomal DNA from *Xenopus laevis*. *J. Mol. Biol.* **118:** 27–47.

Bird A.P and Wolffe A.P. 1999. Methylation-induced repression—Belts, braces and chromatin. *Cell* **99:** 451–454.

Bird A., Taggart M., Frommer M., Miller O.J., and Macleod D. 1985. A fraction of the mouse genome that is derived from islands of non-methylated, CpG-rich DNA. *Cell* **40:** 91–99.

Bourc'his D. and Bestor T.H. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431:** 96–99.

Boyes J. and Bird A. 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG-binding protein. *Cell* **64:** 1123–1134.

Brandeis M., Frank D., Keshet I., Siegried Z., Mendelsohn M., Nemes A., Temper V., Razin A., and Cedar H. 1994. Sp1 elements protect a CpG island from de novo methylation. *Nature* **371:** 435–438.

Bruniquel D. and Schwartz R.H. 2003. Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat. Immunol.* **4:** 235–240.

Buschhausen G., Wittig B., Graessmann M., and Graessmann A. 1987. Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc. Natl. Acad. Sci.* **84:** 1177–1181.

Chen R.Z., Akbarian S., Tudor M., and Jaenisch R. 2001. Deficiency of

methyl-CpG-binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat. Genet.* **27:** 327–331.

Chen W.G., Chang Q., Lin Y., Meissner A., West A.E., Griffith E.C., Jaenisch R., and Greenberg M.E. 2003. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302:** 885–889.

Cooney C.A., Dave A.A., and Wolff G.L. 2002. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J. Nutr.* **132:** 2393S–2400S.

Cooper D.N. and Youssoufian H. 1988. The CpG dinucleotide and human genetic disease. *Hum. Genet.* **78:** 151–155.

Cooper D.N., Taggart M.H., and Bird A.P. 1983. Unmethylated domains in vertebrate DNA. *Nucleic Acids Res.* **11:** 647–658.

Dennis K., Fan T., Geiman T., Yan Q., and Muegge K. 2001. Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.* **15:** 2940–2944.

Dodge J.E., Okano M., Dick F., Tsujimoto N., Chen T., Wang S., Ueda Y., Dyson N., and Li E. 2005. Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *J. Biol. Chem.* **280:** 17986–17991.

Egger G., Liang G., Aparicio A., and Jones P.A. 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429:** 457–463.

Ehrlich M. 1982. Amount and distribution of 5-methycytosine in human DNA from different types of tissues or cells. *Nucleic Acids Res.* **10:** 2709–2721.

———. 2003. The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clin. Immunol.* **109:** 17–28.

Frank D., Keshet I., Shani M., Levine A., Razin A., and Cedar H. 1991. Demethylation of CpG islands in embryonic cells. *Nature* **351:** 239–241.

Frommer M., McDonald L.E., Millar D.S., Collis C.M., Watt J., Grigg G.W., Molloy P.L., and Paul C.L. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci.* **89:** 1827–1831.

Gaudet F., Hodgson J.G., Eden A., Jackson-Grusby L., Dausman J., Gray J.W., Leonhardt H., and Jaenisch R. 2003. Induction of tumors in mice by genomic hypomethylation. *Science* **300:** 489–492.

Gibbons R.J., McDowell T.L., Raman S., O'Rourke D.M., Garrick D., Ayyub H., and Higgs D.R. 2000. Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat. Genet.* **24:** 368–371.

Guy J., Hendrich B., Holmes M., Martin J.E., and Bird A. 2001. A mouse *Mecp2*-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat. Genet.* **27:** 322–326.

Harbers K., Schnieke H., Stuhlmann H., Jahner D., and Jaenisch B. 1981. DNA methylation and gene expression; endogenous retroviral genome becomes infectious after molecular cloning. *Proc. Natl. Acad. Sci.* **78:** 7609–7613.

Hata K., Okano M., Lei H., and Li E. 2002. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129:** 1983–1993.

Hendrich B., Guy J., Ramsahoye B., Wilson V.A., and Bird A. 2001. Closely related proteins Mbd2 and Mbd3 play distinctive but interacting roles in mouse development. *Genes Dev.* **15:** 710–723.

Hendrich B., Hardeland U., Ng H.-H., Jiricny J., and Bird A. 1999. The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401:** 301–304.

Herman J.G., Graff J.R., Myohanen S., Nelkin B.D., and Baylin S.B. 1996. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci.* **93:** 9821–9826.

Holliday R. and Pugh J.E. 1975. DNA modification mechanisms and gene activity during development. *Science* **186:** 226–232.

Hutchins A., Mullen A., Lee H., Barner K., High F., Hendrich B., Bird A., and Reiner S. 2002. Gene silencing quantitatively controls the function of a developmental *trans*-activator. *Mol. Cell* **10:** 81–91.

Jackson J.P., Lindroth A.M., Cao X., and Jacobsen S.E. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416:** 556–560.

Jahner D., Stuhlmann H., Stewart C.L., Harbers K., Lohler J., Simon I., and Jaenisch R. 1982. De novo methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* **298:** 623–628.

Jeddeloh J.A., Stokes T.L., and Richards E.J. 1999. Maintenance of genomic methylation requires a SW12/SNF2-like protein. *Nat. Genet.* **22:** 94–97.

Jones P.A. and Taylor S.M. 1980. Cellular differentiation, cytidine analogues and DNA methylation. *Cell* **20:** 85–93.

Jones P.L., Veenstra G.J., Wade P.A., Vermaak D., Kass S.U., Landsberger N., Strouboulis J., and Wolffe A.P. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat. Genet.* **19:** 187–191.

Kaneda M., Sado T., Hata K., Okano M., Tsujimoto N., Li E., and Sasaki H. 2004. Role of de novo DNA methyltransferases in initiation of genomic imprinting and X-chromosome inactivation. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 125–129.

Kass S.U., Landsberger N., and Wolffe A.P. 1997. DNA methylation directs a time-dependent repression of transcription initiation. *Curr. Biol.* **7:** 157–165.

Kawasaki H. and Taira K. 2004. Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. *Nature* **431:** 211–217.

Klose R.J., Sarraf S.A., Schmiedeberg L., McDermott S.M., Stancheva I., and Bird A.P. 2005. DNA binding specificity of MeCP2 due to a requirement for A/T sequences adjacent to methyl-CpG. *Mol. Cell* **19:** 667–678.

Lehnertz B., Ueda Y., Derijck A.A., Braunschweig U., Perez-Burgos L., Kubicek S., Chen T., Li E., Jenuwein T., and Peters A.H. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13:** 1192–1200.

Lei H., Oh S.P., Okano M., Juttermann R., Gos K.A., Jaenisch R., and Li E. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122:** 3195–3205.

Li E., Bestor T.H., and Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915–926.

Lin I.G., Tomzynski T.J., Ou Q., and Hsieh C.L. 2000. Modulation of DNA binding protein affinity directly affects target site demethylation. *Mol. Cell. Biol.* **20:** 2343–2349.

Lindahl T. 1974. An *N*-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl. Acad. Sci.* **71:** 3649–3653.

Luikenhuis S., Giacometti E., Beard C.F., and Jaenisch R. 2004. Expression of MeCP2 in postmitotic neurons rescues Rett syndrome in mice. *Proc. Natl. Acad. Sci.* **101:** 6033–6038.

MacLeod D., Charlton J., Mullins J., and Bird A.P. 1994. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev.* **8:** 2282–2292.

Martienssen R.A., Doerge R.W., and Colot V. 2005. Epigenomic mapping in *Arabidopsis* using tiling microarrays. *Chromosome Res.* **13:** 299–308.

Martinowich K., Hattori D., Wu H., Fouse S., He F., Hu Y., Fan G., and Sun Y.E. 2003. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302:** 890–893.

Mayer W., Niveleau A., Walter J., Fundele R., and Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature* **403:** 501–502.

McKeon C., Ohkubo H., Pastan I., and de Crombrugghe B. 1982. Unusual methylation pattern of the alpha 2(I) collagen gene. *Cell* **29:** 203–210.

Meehan R.R., Lewis J.D., McKay S., Kleiner E.L., and Bird A.P. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58:** 499–507.

Millar C.B., Guy J., Sansom O.J., Selfridge J., MacDougall E., Hendrich B., Keightley P. D., Bishop S.M., Clarke A.R., and Bird A. 2002. Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science* **297:** 403–405.

Mohandas T., Sparkes R.S., and Shapiro L.J. 1981. Reactivation of an inactive human X-chromosome: Evidence for X-inactivation by DNA methylation. *Science* **211:** 393–396.

Nan X., Ng H.-H., Johnson C.A., Laherty C.D., Turner B.M., Eisenman R.N., and Bird A. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393:** 386–389.

Okano M., Xie S., and Li E. 1998. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **19:** 219–220.

Okano M., Bell D.W., Haber D.A., and Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99:** 247–257.

Petronis A. 2001. Human morbid genetics revisited: Relevance of epigenetics. *Trends Genet.* **17:** 142–146.

Posfai J., Bhagwat A.S., Posfai G., and Roberts R.J. 1989. Predictive motifs derived from cytosine methyltransferases. *Nucleic Acids Res.* **17:** 2421–2435.

Prokhortchouk A., Hendrich B., Jorgensen H., Ruzov A., Wilm M., Georgiev G., Bird A., and Prokhortchouk E. 2001. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev.* **15:** 1613–1618.

Riggs A.D. 1975. X-inactivation, differentiation and DNA methylation. *Cytogenet. Cell Genet.* **14:** 9–25.

Rougier N., Bourc'his D., Gomes D.M., Niveleau A., Plachot M., Paldi A., and Viegas-Pequignot E. 1998. Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12:** 2108–2113.

Sarraf S.A. and Stancheva I. 2004. Methyl-CpG-binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Mol. Cell* **15:** 595–605.

Simonsson S. and Gurdon J. 2004. DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nat. Cell Biol.* **6:** 984–990.

Stein R., Razin A., and Cedar H. 1982. In vitro methylation of the hamster adenine phosphorybosy transferase gene inhibits its expression in mouse L cells. *Proc. Natl. Acad. Sci.* **79:** 4418–3422.

Stewart C.L., Stuhlmann H., Jahner D., and Jaenisch R. 1982. De novo methylation and infectivity of retroviral genomes introduced into embryonal carcinoma cells. *Proc. Natl. Acad. Sci.* **79:** 4098–4102.

Tamaru H. and Selker E.U. 2001. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414:** 277–283.

Tufarelli C., Stanley J.A., Garrick D., Sharpe J.A., Ayyub H., Wood W.G., and Higgs D.R. 2003. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat. Genet.* **34:** 157–165.

Vardimon L., Kressmann A., Cedar H., Maechler M., and Doerfler W. 1982. Expression of a cloned adenovirus gene is inhibited by in vitro methylation. *Proc. Natl. Acad. Sci.* **79:** 1073–1077.

Venolia L., Gartler S.M., Wasserman E.R., Yen P., Mohandas T., and Shapiro L.J. 1982. Transformation with DNA from 5 azacytidine-reactivated X chromosomes. *Proc. Natl. Acad. Sci.* **79:** 2352–2354.

Vire E., Brenner C., Deplus R., Blanchon L., Fraga M., Didelot C., Morey L., Van Eynde A., Bernard D., Vanderwinden J.M., et al. 2006. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439:** 871–874.

Wade P.A., Gegonne A., Jones P.L., Ballestar E., Aubry F., and Wolffe A.P. 1999. Mi-2 complex couples DNA methylation to chromatin remodeling and histone deacetylation. *Nat. Genet.* **23:** 62–66.

Walsh C.P., Chaillet J.R., and Bestor T.H. 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat. Genet.* **20:** 116–117.

Waterland R.A. and Jirtle R.L. 2003. Transposable elements: Targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* **23:** 5293–5300.

Watt F. and Molloy P.L. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus late promoter. *Genes Dev.* **2:** 1136–1143.

Weaver I.C., Cervoni N., Champagne F.A., D'Alessio A.C., Sharma S., Seckl J.R., Dymov S., Szyf M., and Meaney M.J. 2004. Epigenetic programming by maternal behavior. *Nat. Neurosci.* **7:** 847–854.

Weber M., Davies J.J., Wittig D., Oakeley E.J., Haase M., Lam W.L., and Schubeler D. 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37:** 853–862.

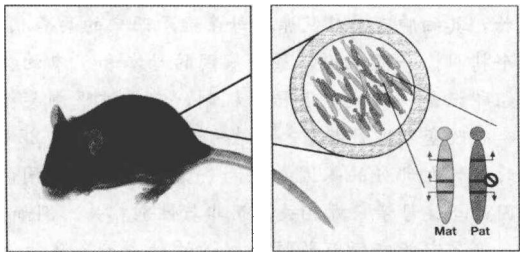
Wigler M.H. 1981. The inheritance of methylation patterns in vertebrates. *Cell* **24:** 285–286.

Wolf S.F., Jolly D.J., Lunnen K.D., Friedman T., and Migeon B.R. 1984. Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X-chromosome: Implications for X-chromosome inactivation. *Proc. Natl. Acad. Sci.* **81:** 2806–2810.

Yoon H.G., Chan D.W., Reynolds A.B., Qin J., and Wong J. 2003. N-CoR mediates DNA methylation-dependent repression through a methyl CpG-binding protein Kaiso. *Mol. Cell* **12:** 723–734.

Zilberman D., Cao X., and Jacobsen S.E. 2003. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299:** 716–719.

# 第 19 章




# 哺乳动物基因组印记

Denise P. Barlow[1] and Marisa S. Bartolomei[2]

[1]*CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Institute of Genetics, Max Perutz Laboratories, A-1030 Vienna, Austria*

[2]*Department of Cell Developmental Biology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6148*

## 本章目录

1. 历史回顾
2. 基因组印记——一种表观遗传的基因调控系统
3. 基因组印记中的关键发现
   - 3.1 印记基因对胚胎和新生儿生长的控制作用
   - 3.2 哺乳动物中基因组印记的功能
   - 3.3 印记基因成簇存在并受印记控制元件的调控
   - 3.4 印记基因簇中至少含有一个非编码 RNA
   - 3.5 DNA 甲基化在基因组印记中的作用
   - 3.6 印记基因簇中两种类型的顺式沉默作用
4. 基因组印记——哺乳动物的一种表观遗传调控模型
5. 展望

致谢

参考文献

## 概要

哺乳动物是二倍体生物，其细胞有着从父亲和母亲继承而来的两套相配对的染色体。因此，哺乳动物的每个基因都有两个拷贝。一般来说，每个基因的母本和父本的拷贝在所有细胞中都有着相同的被激活潜质。基因组印记是一种表观遗传的机制，它通过限制某个基因只能在两条来自父母亲染色体中的一条表达，而改变了这种潜质。在我们整个基因组的 25 000 个基因中，只有几百个基因会产生这个现象，此外的大部分的基因都在两条染色体中有相同的表达方式。雄性和雌性都受基因组印记的影响，因此与父母亲来源相关，而非与性别相关。例如，一个印记基因如果在母源染色体中是活跃表达的，那么它能在所有雄性和雌性的母源染色体上激活，而在所有雄性和雌性的父源染色体上沉默。

在这里，基因组印记的定义被限定于“二倍体细胞中的父母亲特异的基因表达”。因此，二倍体细胞含有所有基因的两个父母亲拷贝，但只表达某个印记基因的一个拷贝，而另一个拷贝被沉默。与此相反，二倍体细胞中，非印记基因的父源和母源拷贝都会表达。在理解基因组印记的概念时，很重要的一点是区分印记基因和某些其他有明显父母源特异性表达的基因，这些特异性表达却是由父母亲对胚胎不等的遗传物质贡献而造成。不等的父母本遗传物质贡献的例子包括：Y 染色体连锁基因只在雄性中出现，雌性个体中具有逃脱了 X 染色体失活的基因，线粒体基因主要由母亲一方贡献以及一些 mRNA 和蛋白质只存在于精子或卵子基质中。

哺乳动物的许多基因组印记特征都使其成为后基因组学时代的一个神秘的生物学问题。例如，印记为亲本冲突理论中的一个可能的进化给出了相应线索，为母亲对内部生殖系统的适应提供了依据，但也可能仅仅是哺乳动物基因组抵御外来 DNA 序列入侵的一种方式。基因组印记可以被认为是一个挑战智力的现象，主要是由于它提出了这样一个问题——为何二倍体生物会进化出一套沉默系统去放弃二倍体状态的优势；最引人注目的事实在于基因组印记系统中的这些基因大多数编码的是调控胚胎以及新生儿生长的因子。因此，进化的基因组印记很有可能在哺乳动物生殖中是起到特定作用的。

在我们理解的层次上，基因组印记似乎没有在原生动物、真菌、植物和动物这四种真核生物的王国中广泛存在。然而它确实以一种很可能的形式存在于两类无脊椎节肢动物——球虫（Coccidae）和眼蕈蚊（Sciaridae）以及一些种子植物，如玉米和拟南芥的胚乳中。这种分布暗示了基因组印记在生命进化中至少三次独立产生。令人惊奇的是，尽管这暗示了基因组印记有独立的进化，但这些印记机制间却表现出一些相似之处。这可能反映了基因组印记和正常基因调控都依赖的基础表观遗传调节机制具有保守性。

## 1. 历史回顾

哺乳动物的基因组印记在医学、社会学以及知识层面都有着极大的意义：①临床上对遗传特征和疾病的控制；②使用辅助生育技术控制人类和动物繁育的能力；③生物技术和后基因组医疗研究的进步。如今讨论任何遗传问题，不管是研究还是医疗，都有必要考虑某个基因是否是父母双源（二倍体）表达或受到基因组印记的影响而父母单源特异性（单倍体）的表达。令人惊讶的是，尽管基因组印记对人类健康有如此重要的作用，它的存在却直到 20 世纪 90 年代初期才被广泛接受，在这之前发现小鼠中有三个基因明确地具有父母源特异性表达。

整个染色体的父母源特异性行为早在 20 世纪 30 年代就在节肢动物染色体的细胞遗传学研究中得以观察（Chandra and Nanjundiah 1990）。有趣的是，“染色体印记”这一词语是在描述一些三节肢动物种属性别决定中起重要作用的父源特异性染色体消除现象中被首次用到（Crouse et al. 1971）。值得注意的是，哺乳动物 X 染色体的染色体印记可导致所有雌性有袋动物细胞及老鼠胚外组织内的一条 X 染色体发生父源特异性染色体失活（Cooper et al. 1971）。与此同时，经典遗传学家们正在做带有染色体易位突变小鼠，这为观察到印记基因表达提供了条件。一些本来被用作定位染色体上基因的“易位”小鼠显示出父母源特异性的表型。它们有两份来自亲本一方的特定染色体区域，而没有从另一方得到这被称作单亲双体型（uniparental disomy，UPD）（图 19-1）。这些结果暗示了“特定的父源或母源基因的单倍表达对于小鼠正常发育很重要”（Searle and Beechey 1978）这一可能性。同时，另一些遗传学家用一种特殊的小鼠突变体，即在 17 号染色体上有大段缺失的“发卡-尾”（Hairpin-Tail）小鼠，明确推翻了遗传学的一个基本命题“在给定位点杂合的生物体表型相同，不管是哪个配子贡献了哪个等位基因给这个基因型”（Johnson，1974）。相反，从母亲得到“发卡-尾缺失”小鼠体积增大，而且在胚胎发育的中途就会死亡；而从父源传递得来遗传学上相同的染色体的小鼠则可以存活并且可育（图 19-1）。事后注意到，对于这些遗传易位和缺失的实验，尽管在哺乳动物中有印记 X 染色体失活的现象，但更被认同的解释是这些常染色体上的基因主要在单倍体的卵子或精子中起作用并修饰后来胚胎发育过程中用到的蛋白质。尽管如此，父源和母源基因组差异化功能的概念渐渐得以站稳脚跟，并有“母源基因组的”发卡-尾“染色体区域中在正常情况下可能是活跃的，而父源的相同部分却被特异地失活了”（McLaren 1979）这种说法的提出。

在证明哺乳动物基因组印记存在的过程中，一个重要的进步是几年后使用核移植技术来试验能否只用小鼠卵子细胞核来产生二倍体单亲胚胎。核移植技术从一个刚刚受精的受精卵中取出一个供体雌原核或者雄原核，并用精密的微注射管将其置于受体受精卵（已经相应的去除了雄原核或雌原核）中，产生新的二倍体胚胎，不同的是该胚胎含有两套母源或父源基因组（分别被叫做孤雌和孤雄胚胎，gynogenetic and androgenetic embryos）（图 19-2）。这项技术首先被用来证明受精的“发卡-尾”突变鼠胚胎中取得的细胞核转移到野生型受体卵子中并不能存活。这证明了，不是卵子的细胞质，而是胚

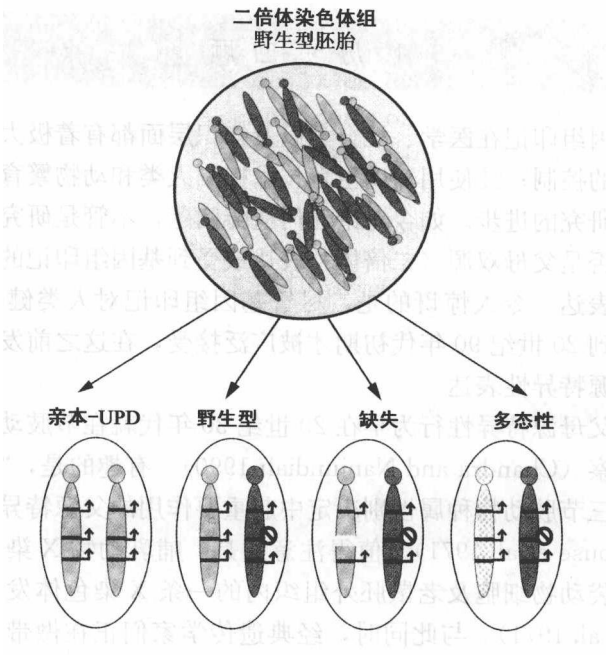



图 19-1 用于区分父母源染色体基因组印记研究的小鼠模型

哺乳动物是二倍体，从双亲中各继承一套完整的染色体，然而小鼠可以产生①一对染色体的两个拷贝都来自于亲本一方，从另一方没有得到任何拷贝②从亲本一方得到一个部分缺失的染色体，而从另一方得到野生型染色体③从一方得到有单核苷酸多态性的染色体而从另一方得到野生型。UPD 或缺失染色体的小鼠很可能有致死的表型，而 SNP 能产生可存活的子代。

胎基因组带有发卡-尾的缺陷；还确定了这样的设想——在胚胎发育中 17 号染色体上基因的父源和母源拷贝存在功能差异（McGrath and Solter 1984b）。后来，人们又用核移植来表明由两个雌原核重建的胚胎（孤雌胚胎）不能存活；只有由一个雌原核一个雄原核重建的胚胎才能产生可存活、可育的子代（McGrath and Solter 1984a；Surani et al. 1984）。这项研究推翻了之前关于单亲小鼠可以活到成年（Hoppe and Illmensee 1982）的结论。孤雌生殖胚胎在死亡时是缺少贡献胎盘的胚胎外组织，而孤雄胚胎则是胚胎组织有缺陷。由此产生了一种假说，胚胎发育需要母源基因组中印记基因的表达而父源基因组表达了胚胎外发育相关的印记基因（Barton et al. 1984）。此后，小鼠中并没有发现印记基因必须有以上的特点，所以孤雌和孤雄胚胎中观察到的差异可能由某一个或某一些印记基因的主导效应来解释。

核移植实验和小鼠遗传学中支持性数据的结合，令人信服地证明了双亲的基因组在小鼠胚胎发育中都是必需的，也为证明哺乳动物基因组印记的存在奠定了坚实的基础（图 19-2）。一项研究使用了“易位”小鼠产生 UPD 染色体（图 19-1），进而深入探求亲本染色体对胚胎发育的贡献，该实验发现，小鼠 2 号、11 号染色体上两个区域是两个母源拷贝或两个父源拷贝时会产生相反的表型。这进一步证明了哺乳动物有亲本特异性基因表达的论断（Cattanach and Kirk 1985）。另外人类临床遗传学得到的数据也正在收集中，这些能强力地证明一些遗传现象的最佳的解释是亲本特异性基因表达，其中

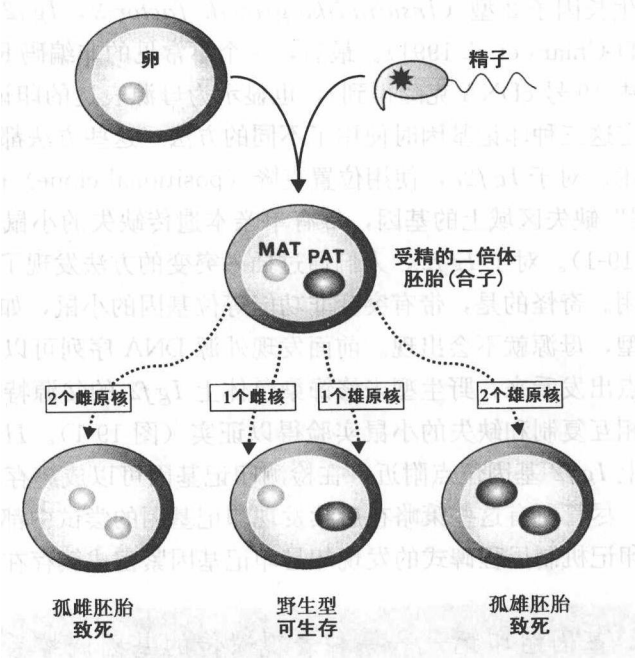



图 19-2　哺乳动物繁殖中需要一套父源和一套母源基因组

核移植技术使用微注射管和高倍显微镜从刚受精的卵子中取出雌（或雄）原核并将其以不同组合方式植入另一个已经被去核“宿主”受精卵，产生新的二倍体胚胎，它们具有两套母源（孤雌）或父源（孤雄）基因组，或有双亲基因组。孤雌和孤雄胚胎在胚胎发育早期就死亡，只有含一个雌核和一个雄核的重建胚胎（野生型）可以存活并繁殖后代。这些实验表明在哺乳动物生殖中母源和父源基因组都是必需的，而且暗示两套亲本基因组表达了完整胚胎发育所需要的不同的一整套基因。

最值得注意的例子有 Prader-Willi 症（似乎专由父源遗传引起）等（Reik 1989）。新技术（将特定基因序列显微注射到小鼠受精卵中）产生的转基因小鼠也可以提供进一步的依据，但这个技术经常被 DNA 甲基化的问题所困扰，因为这种甲基化常常出人意料地引起体细胞组织中转基因的沉默。然而这个“问题”本身就更为亲本染色体有不同功能这一观点增加了筹码，因为有实验表明一些转基因在它们获得 DNA 甲基化的能力上也表现出亲本特异的差异。这种特异性常常表现为母源遗传的转基因得到甲基化而父源的没有，尽管在一小部分情况下 DNA 甲基化差异与父源特异性表达相对应。虽然之后发现了许多“转基因”甲基化印记与小鼠内源基因组印记之间的多处相似，但它们之间有一些可以区分的特征（Reik et al. 1990）。这包括对背景效应的高度易感性，即在很多情况下，印记行为需要一种混合背景才显露；不能在不同染色体整合位点维持印记的表达，以及表现印记效应时对外源 DNA 序列的需要（Chaillet et al. 1995）。

尽管有很多支持性的数据，哺乳动物存在基因组印记的最终证据要靠发现显示亲本特异性表达基因。1991 年，三个小鼠印记基因终于发现。其中第一个，类胰岛素生长因子 2 型受体（*Insulin-like growth factor receptor-2*，*Igf2r*，是生长激素基因 *Igf2* 的“清道夫”受体）是母源表达的印记基因。后来发现这个基因可以解释“发卡-尾”小鼠突变体过度生长的表型（Barlow et al. 1991）。几个月后，已知具有生长激素功能

的基因类胰岛素生长因子 2 型（*Insulin-like growth factor-2*，*Igf2*）被发现是父源表达的印记基因（DeChiara et al. 1991）。最后，一个不常见的非编码 RNA *H19* 基因（从胎儿肝细胞文库中 19 号 cDNA 克隆得到），也显示为母源表达的印记基因（Bartolomei et al. 1991）。确定这三种印记基因时使用了不同的方法，这些方法都用到了小鼠遗传学中新兴的实验技术。对于 *Igf2r*，使用位置克隆（positional clone）确定了位于 17 号染色体上“发卡-尾”缺失区域上的基因，带有单亲本遗传缺失的小鼠用于发现母源特异表达的基因（图 19-1）。对于 *Igf2*，人们通过插入突变的方法发现了其在胚胎发育中生长因子的生理作用。奇怪的是，带有突变非功能等位基因的小鼠，如果基因来源于父源才会出现相应表型，母源就不会出现。前面发现外源 DNA 序列可以引起小鼠基因的印记表达，从此观点出发看来，野生型未修饰染色体上 *Igf2* 的父源特异性表达也从 7 号染色体上亲本的相互复制和缺失的小鼠实验得以证实（图 19-1）。*H19* 非编码 RNA 定位于 7 号染色体上 *Igf2* 基因位点附近，在检测印记基因可以成簇存在的假说时，发现它也是印记基因。尽管所有这些策略在后续发现印记基因的尝试中都是有用的，但理解哺乳动物基因组印记机制里程碑式的发现却是印记基因紧密成簇存在的事实。

## 2. 基因组印记——一种表观遗传的基因调控系统

顺式作用是基因组印记的决定性特征（见 Box 1）。因此印记机制只在一条染色体上起作用；相反，反式基因调控机制可以在细胞核中任何染色体上起作用。如果是互交种群，两条亲本染色体一般只含有许多单碱基对差异，这被称为单核苷酸多态性（single nucleotide polymorphism，SNP），但如果使用自交小鼠系，它们应该在遗传背景上是相同的。因为在含有相同亲本染色体的自交小鼠中有基因组印记的存在，此过程必定使用了一种表观遗传学机制来修饰 DNA 序列所带有的信息并且制造出两个亲本基因拷贝的差异。这些发现也暗示存在一种只限于一条染色体的顺式沉默机制，因此沉默因子不能在核中自由散布到达活跃的基因拷贝。尽管印记基因在一条亲本染色体上受抑制而在另一条上活跃，但我们不知道基因组印记是否只是一个沉默机制，也许相反，我们必须考虑到它可能是激活了哺乳动物基因组中本来应该沉默基因。

Box 1 哺乳动物基因组印记的关键特征

- 顺式作用机制
- 是继承而非性别的结果
- 印记是一个亲本配子得到的表观遗传修饰
- 印记基因大多和一个非编码 RNA 聚集在一起
- 印记可以修饰作用于多种基因的长程调控元件
- 印记基因在哺乳动物发育中起到作用

因此基因组印记的起点必须依赖一种修饰或“印记”一条亲本染色体的表观遗传系统（图 19-3）。我们可以推论，这个印记随后就可以吸引或排斥转录因子然后改变一条

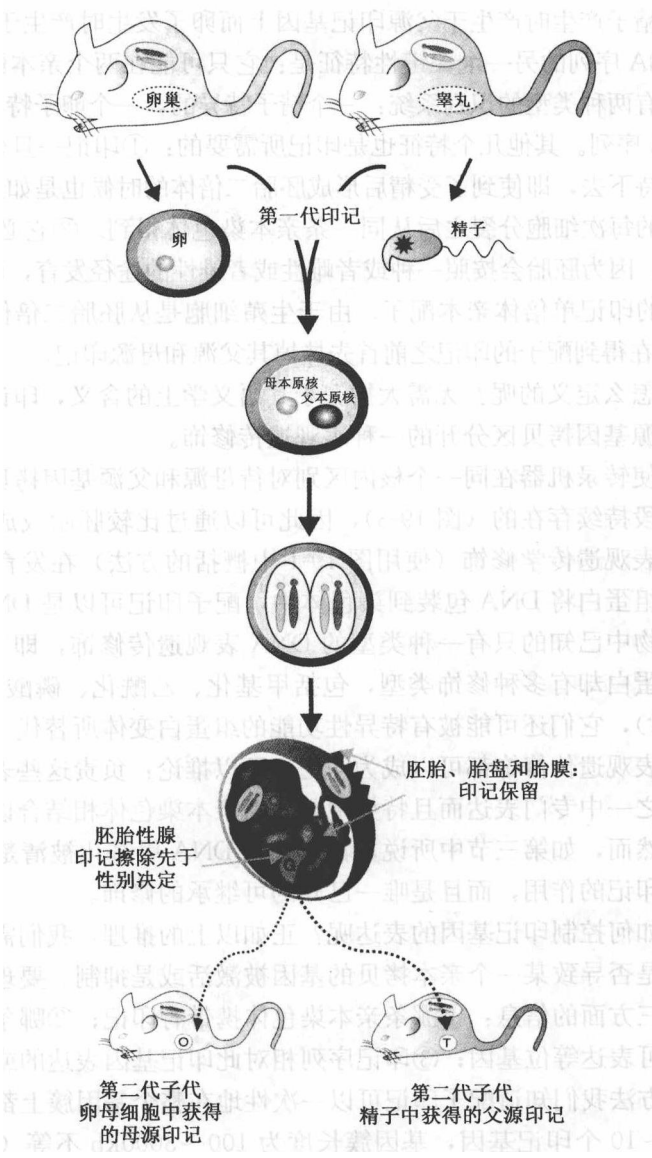



图 19-3　哺乳动物发育中的印记获得和擦除

印记是在配子中建立的，因此卵子和精子已经携带了有印记的染色体（第一代印记）。受精后胚胎成为二倍体，胚胎、膜、胎盘和成体中细胞不断分裂，印记依然保持在相同的亲本染色体上。生殖细胞是在胚胎性腺中形成的，仅在这些细胞中，印记会在性别决定之前被擦除。当胚胎发育成雄性，性腺分化成睾丸，产生单倍体精子，它们在染色体上获得了父源印记。类似的，在雌性发育时，卵子中染色体得到母源印记（第二代印记）。

亲本染色体上基因的表达。已知遗传学上染色体相同的自交小鼠也显示有基因组印记，可以推理亲本印记不可能在胚胎形成二倍体之后获得，因为细胞的表观遗传系统无法区分两个完全相同的亲本基因拷贝。因此亲本印记必须是在两套亲本染色体分开的时候发

生的，而这只有在配子形成过程和受精后约 12 小时之内发生（图 19-3）。最有可能的情况是配子印记在精子产生时产生于父源印记基因上而卵子发生时产生于母源印记基因上。关于“印记”DNA 序列的另一个关键性特征是：它只可能在两个亲本配子中的一个中被修饰；因此需要有两种类型的识别系统：一个精子特异的，一个卵子特异的，二者都靶向一段不同的 DNA 序列。其他几个特征也是印记所需要的：①印记一旦建立就必须在相同亲本染色体上维持下去，即使到了受精后形成胚胎二倍体的时候也是如此。②印记必须在胚胎及成体动物的每次细胞分裂之后从同一条亲本染色体得到。③它必须是可以擦除的。最后一点很重要，因为胚胎会按照一种或者雌性或者雄性的途径发育，而它们的性腺需要只产生一种类型的印记单倍体亲本配子。由于生殖细胞是从胚胎二倍体细胞而来的（图 19-3），它们必须在得到配子的印记之前首先抹掉其父源和母源印记。

配子印记是怎么定义的呢？无需太拘泥于其语义学上的含义，印记的概念只是将母源基因拷贝从父源基因拷贝区分开的一种表观遗传修饰。

印记必须能使转录机器在同一个核内区别对待母源和父源基因拷贝。预计配子印记是在所有发育阶段持续存在的（图 19-3），因此可以通过比较胚胎或成体组织的母源和父源染色体上的表观遗传学修饰（使用图 19-1 中概括的方法）在发育中追溯到它的来源而找到印记。组蛋白将 DNA 包装到染色体中，配子印记可以是 DNA 或组蛋白的修饰。尽管哺乳动物中已知的只有一种类型的 DNA 表观遗传修饰，即 DNA 甲基化（见第 18 章），但组蛋白却有多种修饰类型，包括甲基化、乙酰化、磷酸化、sumo 化及泛素化（见第 10 章），它们还可能被有特异性功能的组蛋白变体所替代（见第 3、13 章）。理论上任何这些表观遗传修饰都可以成为印记。可以推论：负责这些表观遗传修饰的酶类会在两个配子之一中专门表达而且特异地与一条亲本染色体相结合以便在细胞分裂时复制这些修饰。然而，如第三节中所说，目前只有 DNA 甲基化被清楚地证明在哺乳动物中起到了配子印记的作用，而且是唯一已知的可继承的修饰。

配子印记是如何控制印记基因的表达呢？正如以上的推理，我们需要“保持开放的想法”来看印记是否导致某一个亲本拷贝的基因被激活或是抑制。要理解印记是如何运作的，我们需要三方面的信息：①哪条亲本染色体携带有印记；②哪条亲本染色体带有这个印记基因的可表达等位基因；③印记序列相对此印记基因表达的或沉默的等位基因的位置。用这种方法我们知道配子印记可以一次性地在整个基因簇上都起作用。这些印记基因簇含有 3～10 个印记基因，基因簇长度为 100～3000kb 不等（详细信息请登录 http://www.mgu.har.ac.uk/research/imprinting/）。印记基因簇中大多数都是的编码蛋白 mRNA 的印记基因，但至少有一个非编码 RNA 印记（non-coding RNA，ncRNA）基因。

由于印记基因在基因簇中的排列形式是让一些基因从一条亲本染色体上得到表达而另一些从另一条上表达，所以要弄清楚印记如何运作并不总是那么简单。研究印记对基因簇中某一单个基因的作用是可能的，但研究它在整个基因簇的效应会进一步拓宽人们的认识，这将在第 3 节更详细地阐述。有一件事情却是清楚的：大自然没有选取最简单的模式——印记靶向启动子，而是优先沉默配子中的印记基因。相反，印记似乎通常是靶向了在相同染色体上影响多个基因表达的长程顺式作用抑制子。

## 3. 基因组印记中的关键发现

### 3.1　印记基因对胚胎和新生儿生长的控制作用

哺乳动物的基因组印记有什么功能？回答这个问题的方法之一是确定已知印记基因在体内的作用。现代技术使得小鼠基因功能可以通过将基因序列突变功能缺失来确定。运用这种“同源重组”技术，在 78 个抑制印记基因中确定了 26 个基因的功能（原始参考文献见 http://www.mgu.har.ac.uk/research/imprinting/function.html）。表 19-1 列出了这些基因，依据它们在小鼠发育中功能及其表达的亲本等位基因进行分类排列。至今最大的一类包括了影响胚胎、胎盘、或（完全依靠母乳的）新生儿生长的印记基因。在这一大类中约有一半是有生长促进因子作用的父源表达印记基因（胚胎中此基因缺陷表现为生长迟滞）；另一半是有生长抑制因子作用的母源表达印记基因（胚胎中此基因缺陷则表现为生长亢进）。仅次于此一类的基因如缺失也不会引起胚胎发育明显缺陷。第三大类基因若缺失则产生行为或神经上的缺陷。剩下的 3 个被测基因导致多种彼此无关的缺陷。这些结果没有为所有印记基因指明功能，在某种意义上令人失望；然而在黑暗的一端可能有光亮，因为这些结果确实告诉我们超过 50%的印记基因有胚胎或新生儿生长调控因子的作用。更有趣的是，母源表达的生长调控基因会抑制子代生长，而相同分类中父源表达基因却促进生长。所研究的印记基因中有 20%活跃于神经过程中，其中一些通过改变母源行为而影响新生儿生长速率。最令人困惑的是在胚胎发育中没有明显生物学功能的一类，大约占了所测印记基因的 25%，这与试图找到选择压力获得印记基因表达这个出发点不相符。

**表 19-1　印记基因的功能(由基因缺失实验得出)**

| 母源 | 基因功能 | 父源 |
|---|---|---|
| −*Igf2r* | 胚胎、胎盘和出生后阶段的发育缺陷 | +*Igf2* |
| −*Gnas* | | +*Gnasxl* |
| −*Tssc3*/*Ipl* | | +*Pegl*/*Mest* |
| −*Mash2* | | +*Peg3*/*Pwl* |
| −*Grb10*/*Megl* | | +*Rasgrf1* |
| −/+*Cdknlc* | | +*Dlkl* |
| *Nesp* | | +*Pegl*/*Mest* |
| *Ube3a* | 行为或神经缺陷 | +*Peg3*/*Pwl* |
| *Kcnql* * | | +*Rasgrf1* |
| *Asbl* 精子发生 | 其他缺陷 | *Ndn* 品系特异性致死 |
| *Dcn* 肿瘤抑制因子 | | |
| *H19 ncRNA* | | *Snrpn*/*Snurf* |
| *Slc22a2* | 没有胚胎或新生儿的明显缺陷 | *Frat3* |
| *Slc22a3* | | *Ins2* |

注：(母源)母源表达印记基因，(父源)父源表达印记基因，(+)生长促进效应(−)生长抑制效应 (−/+)分化中有缺陷但生长调节状态不明，(*)附加的分化缺陷(原始数据请参考 http://www.mgu.har.mrc.ac.uk/research/imprinting/function.html)。

## 3.2 哺乳动物中基因组印记的功能

基因功能分析能帮助我们理解为何哺乳动物存在基因印记。对不同种类哺乳动物的基因组印记的研究给出了一些思路。胎生哺乳动物如小鼠和人类以及有袋类（如负鼠和小袋鼠）都有基因组印记。卵生哺乳动物（如鸭嘴兽和针宴鼠）看上去缺少印记基因（尽管深入的研究还没有进行）。胎生哺乳动物和有袋类动物有不同生殖方式，但这两种方式都直接影响胚胎对母亲的营养吸收。相反，在卵中发育的胚胎不能直接影响到母源物质。许多脊椎和非脊椎类动物都采取卵生，值得注意的是，它们还可以进行孤雌生殖——一种雌性配子不经过雄性配子的受精而直接发育成一个新的二倍体个体的生殖方式（注意：孤雌胚胎由相同母源基因组复制产生，而图 19-3 中描述的双雌胚胎则由两套不同的母源基因组产生）。生物体进行孤雌生殖的能力很可能暗示了它们完全没有基因组印记，因为这显示了其父源基因组是可以或缺的。然而在哺乳动物中，印记基因表达控制胎儿发育的直接后果是不能进行孤雌生殖，母源和父源亲本对产生可存活的后代所都是必需的，所以哺乳动物完全依靠有性生殖产生可存活子代（图 19-4）。孤雌生殖迄今没有在哺乳动物中发现，尽管有一些相反的提法。但有报道指出通过操纵 *Igf2* 这一印记基因簇，已经产生了一些具有二倍体母源基因组的稀有小鼠（Kono et al. 2004）。

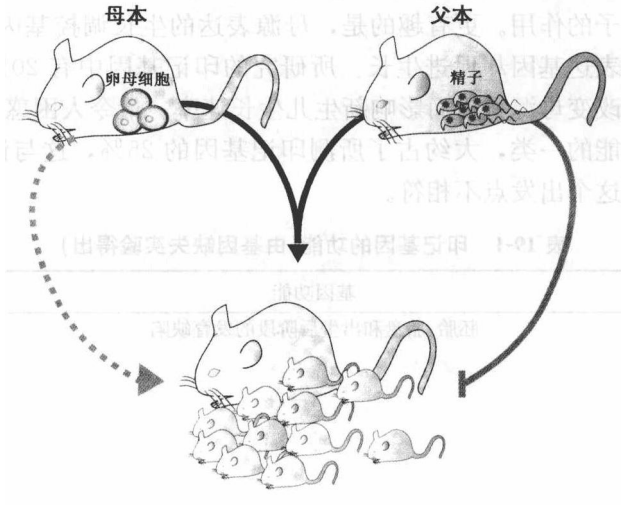



图 19-4 印记基因在哺乳动物生殖中起到作用

哺乳动物是二倍体，单倍体的卵子被单倍体精子结合之后才能产生二倍体胚胎。只有雌性在结构上被赋予了生殖的能力，但是它们不能以孤雌生殖的方式来繁殖，因为胎儿生长所需的重要基因是有印记的而且在母源染色体上被沉默了，这些基因只在父源染色体上表达。因此两套亲本基因组都是哺乳动物生殖所必需的。孤雌生殖：由相同母源基因组的两份拷贝产生二倍体子代。

为何通常基因组印记仅存在于一些哺乳动物中而不从脊椎动物中进化出来呢？这与基因组印记的三大特征有关：即多数印记基因有生长调节功能；印记基因的存在限于胎生和有袋类动物；父源基因对胎儿发育的必需性。此结论支持了下述两个同样吸引人的假说。

第一个假说提出基因组印记是对“亲本冲突”产生的应答而进化来的（Moor and Haig 1991），这种冲突是由于母源、父源基因组有相反的利益所引起。胚胎生长依赖一个亲本但是受含有两个亲本基因组的胚胎的影响。父源表达的印记基因被认为增加胚胎生长，而使含有特定父源基因组的子代个体竞争力最大化；母源表达的印记基因被认为抑制胎儿生长，这使得母源物质能更公平地分给所有子代并且增加了母源基因组向多个子代的传递，但它们却可能含有不同的父源基因组。

第二个假说叫做“滋养层防御”（trophoblast defense）（Varmuza and Mann1994）。认为母源基因组由于结构上能进行内部繁殖，如果卵子自发的激活会引起全面胚胎发育，母体就会受到威胁；而雄性缺乏这种内部繁殖必需的解剖学结构，所以即便精原细胞自然被激活，它也没有威胁。因此认为印记要么沉默了母源染色体上促进胎盘发育的基因，要么激活了限制这一进程的基因。所以胎盘形成必需的基因只能在受精后从父源基因组中得到表达。

如果可能，那么二者之中哪个假说能够正确解释哺乳动物的基因组印记的进化呢？两个假说都提出了印记基因在调控胎盘发育和功能的作用，然而亲本冲突和滋养层防御模型都没有完全解释所有的现象（Wilkins and Haig 2003）。人们注意到一个有趣的现象，就是在植物胚乳中也发现了印记基因。胚乳被比作哺乳动物的胎盘，因为它将营养物质从亲本植物传递到胎盘（见第 11 章）。这个发现增强了以下论点的说服力：基因组印记进化成调节亲体和子代间营养传递的一种方式。更完全或替代性地解释哺乳动物当中基因组印记功能能来自两类实验结果，①在整个基因簇检查“印记”的功能，而不是仅仅使小鼠某单个印记基因缺失观察表型。②准确得到基因是如何被印记的。有可能一个簇中不是所有基因都是印记机制的特定目标而一些可能只是该进程“无辜的旁观者”，所以它们的功能对于基因组印记的作用来说没有提供太多信息。被印记机制影响的“无辜的旁观者”基因的存在令人满意地解释了为何在发育中没有明显生物学功能的印记基因会如此之多（表 19-1）。

### 3.3　印记基因成簇存在并受印记控制元件的调控

迄今发现十条小鼠染色体共有大约 80 个印记基因，大多数以基因簇形式存在（Verona et al. 2003）。8 条染色体上（分别为 2 号、6 号、7 号、9 号、11 号、12 号、15 号和 17 号）共有 11 个印记基因簇分布，而只在 3 条染色体上发现（为 2 号、14 号和 18 号）单个的印记基因。印记基因簇的存在强烈暗示了某一共同的 DNA 元件会顺式调节多个印记基因的表达。目前这些印记基因簇中只有 6 个比较了解。表 19-2 列出了主要印记基因的名字（*Igf2r*、*Igf2*、*Kcnq1*、*Gnas*、*Dlk1* 等基因簇）或疾病相关的 *Pws* 簇，引起 Prader-Willi 综合征，将在第 23 章中详细讨论。这六个基因簇含有 3～10 个印记基因，分别为 100～3000kb 的 DNA 片段。

图 19-5 显示了典型印记基因簇中的亲本特异性表达。这 6 个簇的共同特征是它们都有一个带配子甲基化印记的 DNA 区域，叫做配子差异性 DNA 甲基化区域（differentially DNA-methylated region，DMR）。如果甲基化印记在一个配子中建立而且只在胚胎二倍体细胞的一条亲本染色体上维持，就叫做配子甲基化印记（gametic methyla-

tion imprint)。在四个簇中（*Igf2r*、*Kcnq1*、*Gnas*、*Pws*），配子 DMR 有着在卵子发生中得到的母源甲基化印记，而在 *Igf2* 和 *Dlk1* 两个簇中有精子发生中获得的父源甲基化印记。在所有 6 个例子当中，配子 DMR 都控制整个或部分基因簇的表达且因此被定为簇的印记控制元件（imprinting control element，ICE）（Spahn and Barlow 2003）。

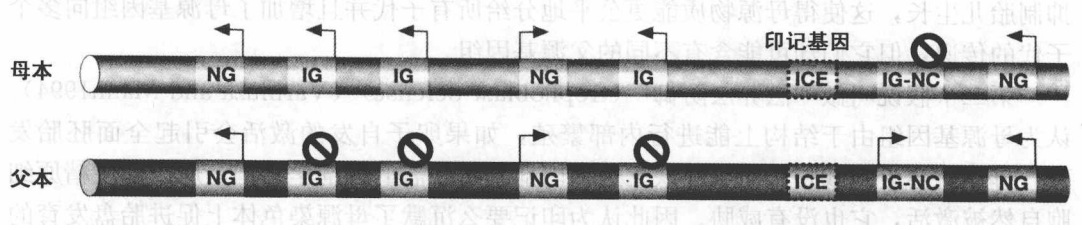



图 19-5 印记基因只在一个亲本等位基因中表达而且常常聚集成簇存在

多数印记基因在基因簇中发现，其中含有多种编码蛋白的 mRNA 和至少一个非编码 RNA（ncRNA）。非印记基因也可能在其中存在。印记机制是顺式作用的，印记表达由含有从一个亲本配子带来的印记控制元件所调控。一对二倍体染色体以粉红色（母源表达的印记基因）和蓝色（父源表达的印记基因）表示。IG，印记的 mRNA 基因；IG-nc，印记的 ncRNA 基因；NG，非印记基因；ICE，印记控制元件；箭头，表达的基因；停止符号，被抑制的基因。

表 19-2 显示出每个印记基因簇都对应多个 mRNA 和至少一个 ncRNA。其中四个基因簇（*Igf2r*、*Kcnq1*、*Igf2* 和 *Dlk1*）的工作方式比较简单，带有甲基化配子 DMR 的染色体表达多个 mRNA，而不表达 ncRNA（如图 19-5 所描述的母源配子 DMR）。携带非甲基化配子 DMR 的染色体表达形式则与前者相反：多个 mRNA 表达被抑制而 ncRNA 得到表达。剩下两个簇（*Gnas* 和 *Pws*）的表达方式很复杂，其印记 mRNA 在两条染色体上都表达，而印记的 ncRNA 只在带有非甲基化 DMR 的染色体上表达。从表 19-2 可以看出，在三个簇（*Igf2r*、*Kcnq1*、*Gnas*）中 ncRNA 启动子是位于印记 mRNA 的一个内含子之中的，而在剩下的簇中，ncRNA 启动子被分开但与印记的 mRNA基因靠得很近。在一个印记基因簇中，紧密混合着活跃和沉默的基因，这暗示着影响印记基因的沉默和激活机制不是分散的而可能限制于被影响基因的附近。尤其是沉默 ncRNA 的启动子位于活跃转录基因的内含子中，这一点暗示着沉默机制甚至可能，仅在调节元件周围，而没有分布在一个基因的全长范围。

配子 DMR 有什么样的作用呢？配子 DMR 可以使母源或父源甲基化，删除这些 DMR 的结果几乎都与此说法一致，但其中也有一些有趣的例外（图 19-6）。在三个簇中（*Igf2r*、*Kcnq1*、*Dlk1*），实验性地删除甲基化配子 DMR 没有影响，然而删除其非甲基化的 DMR 则完全逆转了父源特异性的表达形式并使其失去 ncRNA 表达，产生双等位基因上 mRNA 的同时表达（Lin et al. 1995；Zwart et al. 2001；Fitzpatrick et al. 2002）。有两个簇（*Gnas* 和 *Pws*）所含配子 DMR 似乎不止一个，并且显示出更复杂的行为，但这些行为与这种模式还是有一些相似之处的（Williamson et al. 2006）。*Igf2* 簇则不然：甲基化和非甲基化的 DMR 删除都会引起 mRNA 和 ncRNA 的表达发生顺式变化（Thorvaldsen et al. 1998）。

表 19-2　小鼠基因组印记基因簇特征

| 基因簇类型 | 基因簇名 | 染色体编号 小鼠/人类 | 配子甲基化印记 | 簇大小 (kb) | 簇中基因数量 | mRNA 及其表达 | ncRNA 及其表达 | ncRNA 方向 |
|---|---|---|---|---|---|---|---|---|
| Ⅰ型 | *Igf2r* | 17/6 | M | 400 | 4 | *Igf2r*(M)<br>*Slc22a2*(M)<br>*Slc22a3*(M) | *Air*(P) | 与 *Igf2r* 反义 |
| | *Kcnql* | 7/11 | M | 700 | 10 | *Mash2*(M)<br>*Kcnql*(M)<br>*Cd81*(M)<br>*Cdknlc*(M)<br>*Msuit*(M)<br>*Slc221l*(M)<br>*Ipl*(M)<br>*Tssc4*(M)<br>*Obphl*(M) | *Kcnqlotl*(P) | 与 *Kcnql* 反义 |
| | *Pws* | 7/15 | M | 3000 | 约 7 | *Ube3a*(M) | * *Ube3aas*(P) | 与 *Ube3a* 反义(还与 *Snrpn* 在同义方向上重叠) |
| | | | | | | *Atpl0c*(M)<br>*Frat3*(P)<br>*Mkrn3*(P)<br>*Ndn*(P)<br>*Magel2*(P)<br>*Snrpn*(P) | * *IPW*(P)<br>* *Mkrn3as*(P)<br>* *PEC2*(P)<br>* *PEC3*(P)<br>* *Pwcrl*(P)<br>* 可能为一个长的 *ncRNA* | |
| | *Gnas* | 2/20 | M(×2) | 100 | 5 | *Nesp*(M)<br>*Gnas*(M)<br>*Gnasxl*(P) | [1]Nespas(P)<br>[2]*Exonl A*(P) | [1]与 *Nesp* 反义<br>[2]与 *Gnas* 同义 |
| Ⅱ型 | *Igf2* | 7/11 | P | 100 | 3 | *Igf2*(P)<br>*Ins2*(P) | *H19*(M) | 正义 无重叠 |
| | *Dlk1* | 9/14 | P | 1000 | 7 | *Dlk1*(P)<br>*Dio3*(P)<br>*Rtl1*(P) | *Gtl2*(M) *<br>*Rian*(M) *<br>*Rtl1as*(M) *<br>*Mirg*(M) *<br>* 可能含有更长的 ncRNA | 与 *Dlk1* 同义<br>也和 *Rtl1* 反义 |

注：(M)母源(P)父源 (DMR)差异甲基化区域 细节在文章中给出（经修改摘自 Beechey et al. 2005 [http://www. mgu. har. mrc. ac. uk/research/imprinting])。

以上配子 DMR 删除实验一开始没有暗示到配子 DMR 的共同作用，而要知道其准确功能需要明确 DMR 在每个基因簇中相对于印记基因的位置。在三个有着最简单表达方式的基因簇中（*Igf2r*、*Kcnql*、*Dlk1*）配子 DMR 要么含有要么控制了 ncRNA 的表达；因此缺失此元件会明显导致 ncRNA 表达丢失。*Igf2* 基因簇中的配子 DMR 却不是直接控制 *H19* 启动子的，而是改变了 *Igf2* 和 *H19* 之间的作用以及二者共有的增强子，通过这种方式调控它们的表达。尽管有这些不同，一般来说非甲基化的配子 DMR 在所有六个基因簇中都隐含着有 ncRNA 表达正调控因子的作用，而 DNA 甲基化印记的存在也与 ncRNA 的抑制有关。配子 DMR 删除的实验数据清楚地分辨出这些区域是主要

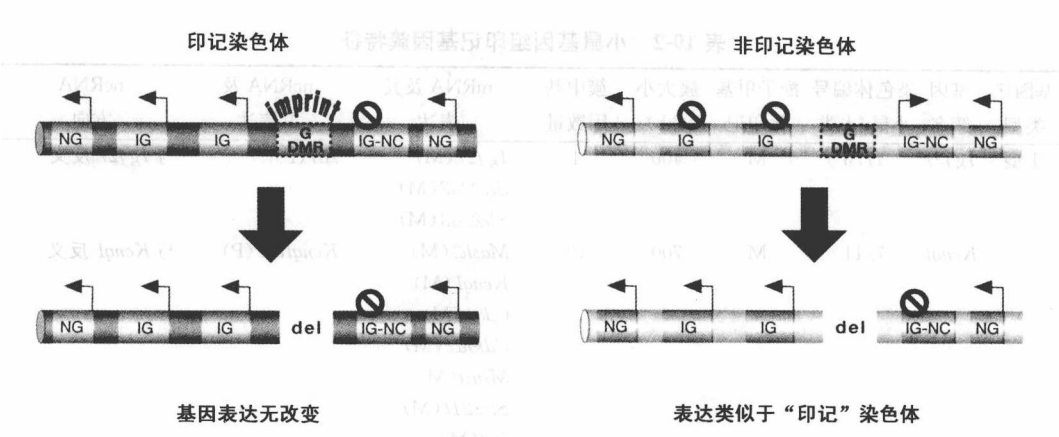



图 19-6 印记表达受配子 DMR 调控

左边为从印记染色体（绿色）删除配子 DMR 的效应。右边为从非印记染色体（黄色）删除了配子 DMR 的效应。在许多印记基因簇中（如 *Igf2r*，*Kcnql* 和 *Dlk1*），实验性删除 G-DMR 只影响带有非印记 G-DMR 的染色体。这导致编码蛋白 mRNA 的印记基因（IG）去抑制而 ncRNA 印记基因（IG-NC）被抑制。注意在此处未标明的一些印记基因簇中（*Igf2* 和 *Pws*），甲基化的 G-DMR 似乎也为顺式表达一些印记 mRNA 所需要。del，被删除的 DNA G-DMR 配子差异甲基化区域；NG，非印记基因；箭头，表达的等位基因；停止符号，被抑制的等位基因；IMPRINT，导致基因表达顺式改变的表观遗传修饰。

的 ICE，其活性被 DNA 甲基化调控。

尽管只有初步数据，人们却已经尝试定义了两种类型的印记基因簇（表 19-2）。Ⅰ型和Ⅱ型基因簇在 ncRNA 和多种 mRNA 之间有颠倒的亲本表达特异性，这个通用性质是相同的，但能被分为三种形式。①Ⅰ型基因簇有着母源 DNA 甲基化印记（如 *Igf2r*、*Kcnql*、*Pws* 和 *Gnas*），而Ⅱ型带有父源 DNA 甲基化印记（*Igf2* 和 *Dlk1*）。②Ⅰ型基因簇从父源染色体上表达 ncRNA 而Ⅱ型基因簇从母源染色体上表达。③Ⅰ型基因簇中 ncRNA 启动子在内含子之中并产生与多个印记 mRNA 基因，并合有着反义方向的转录物，而在Ⅱ型基因簇中启动子在内含子下游并与多个印记 mRNA 基因中至少一个的正义方向相同（Regha et al. 2006）。

### 3.4 印记基因簇中至少含有一个非编码 RNA

在 12 个抑制印记基因簇中有 8 个与 ncRNA 相关（O'Neill 2005）。参与 RNA 加工和翻译过程（如剪接、转移和核糖体 RNA）之外的 ncRNA 之前被认为只占哺乳动物基因组很小一部分。现在，由于已知小鼠和人类的基因组全序列并可以进行转录组学分析，进而可知一个给定细胞群中的所有 RNA 转录物。通过这种分析可知，哺乳动物转录物（不计参与加工和翻译的 ncRNA）中大多数由 ncRNA 组成，而不是编码蛋白的 mRNA，这一点很令人惊讶。这种绝对大量的 ncRNA 转录物存在于哺乳基因组中，而且许多 ncRNA 与已知的编码蛋白的基因重叠，这都暗示着它们不应被视为“转录噪音”，而可能组成一种新的，还未知的基因调控系统（Mattick 2005）。已经有一个名为 NONCODE 的网站收集着所有生物中功能性 ncRNA 的有关数据（http://noncode.bioinfo.org.cn）。有些哺乳动物 ncRNA 有基因调节功能，这类 ncRNA 包括

参与 RNAi 途径的“短”ncRNA（21～30bp）（第 8 章）；“基因之间”调控周围染色质活性的未加工转录产物（Houssecker and Proudfoot 2005）以及“长的”经过加工的 ncRNA，如 X 染色体失活相关的 *Xist*（第 17 章）。

哪些 ncRNA 与印记基因簇相关呢？对表 19-2 中的 6 个印记基因簇相关的 ncRNA 进行的分析虽然还不完全，但得到了一些相似点和差异处。有三个印记 ncRNA 为异常长的成熟 RNA：*Air* 为 108kb（Lyle et al. 2000），*Kcnqlotl* 最终长度还未确定，但至少有 60kb（Mitsuya et al. 1999），而 *Ube3aas* 可能超过 1000kb（Landers et al. 2004）。相反，*H19* ncRNA 只有 2. 3kb（Brannan et al. 1990）。*Gtl2* ncRNA 有多个可变剪接转录产物，而人们也注意到了下游基因间转录，所以还可能有更长的转录单位（Tierling et al. 2005）。*Nespas* ncRNA 很大，甚至无法用 RNA 杂交来分辨，全长仍然未知（Wroe et al. 2000）。所有这些印记的 ncRNA 基因看上去都只含有很少的内含子，即内含子/外显子比率很低，或者像成熟转录物一样不再被剪接。之前有人认为所有印记基因都只含很少内含子，但可能这只对印记的 ncRNA 是正确的（Hurst et al. 1996）。ncRNA 的进一步特征是 *Ube3aas* 和 *Gtl2* 的下游转录物为小核 RNA（small nucleolar RNA，snoRNA，它们可能对 mRNA 进行修饰，是转录后调节因子）和微 RNA（miRNA，转录和转录后 mRNA 调控中涉及）的作用对象。snoRNA 不直接作用于簇中的印记 mRNA 基因，它们可能在印记机制本身没有起到作用（Seitz et al. 2004）。类似的，*Dlk1* 簇中的 miRNA 与簇中一个 mRNA 基因的转录后抑制有关，但还没有检测它们在调控簇的印记表达上的作用（Davis et al. 2005）

印记 ncRNA 的两个特征暗示其对沉默簇中印记 mRNA 基因有作用。①ncRNA 一般在印记 mRNA 基因相反的亲本中特异性表达（表 19-2）；②携带配子甲基化印记的 DMR 控制了整个基因簇的印记表达，三个例子（*Air*、*Kcnqlotl*、*Gnas* 外显子 1a）中，它与 ncRNA 启动子发生重叠。这可能暗示印记是为调节每个印记簇中 ncRNA 进化而来，删除携带配子 DMR 的非甲基化序列的实验印证了这个说法：在实验中，测试 *Igf2r*、*Kcnql*、*Gnas*、*Pws* 和 *Dlk1* 簇（Wutz et al. 1997；Bielinska et al. 2000；Williamson et al. 2006）时发现失去了 ncRNA 表达而获得了印记 mRNA 基因的表达（图 19-6）。*Igf2* 基因簇则是一个例外，因为非甲基化的配子 DMR 似乎不是 *H19* ncRNA 的直接调控子（Thorvaldsen et al. 1998）。

两个来自于Ⅰ型印记基因簇的 ncRNA（*Air* 和 *Kcnqlotl*）和一个来自Ⅱ型簇的 *H19* 被用于实验直接检测 ncRNA 自身的作用。将 108kb 的 *Air* ncRNA 截短到 3kb 后发现 ncRNA 自身对于沉默 *Igf2r* 簇中所有三个 mRNA 基因是必要的，所以这个 ncRNA 有着调节作用（Sleutels et al. 2002）。另外，将 160kb 的 *Kcnqlotl* ncRNA 截短到 1. 5kb 后也发现这个 ncRNA 直接参与沉默更大的 *Kcnql* 簇中所有的十个 mRNA 基因（Mancini-DiNardo 2006）。相反，精确删除了 *H19* ncRNA 及其启动子并没有对内胚层组织中 *Igf2* 簇的印记产生影响，尽管中胚层组织中有一些印记的丢失（Schmidt et al. 1999）。因此两个Ⅰ型母源印记簇有共同 ncRNA 依赖的沉默机制，而这与至今仅有的Ⅱ型父源印记簇使用的是不同的绝缘子依赖模型。哺乳动物是否只有这两种基础印记机制（一个父源制，一个母源）还有待新印记簇的发现。

## 3.5 DNA 甲基化在基因组印记中的作用

1991 年最早发现的三个内源印记基因，使研究者知道了细胞的表观遗传系统如何将一个基因加上亲本印记。最初也是最容易检测的标记为 DNA 甲基化——哺乳动物中共价地给 CpG 中胞嘧啶加上甲基的一种修饰。DNA 甲基化是通过从头甲基转移酶的作用得到的，之后在体内随着细胞分裂，由维持甲基化转移酶来维持（见第 18 章）。因此这种修饰满足图 19-3 中所概括的亲本身份标记或“印记”的标准，因为①它可以在精子或卵子中由只作用于一个配子的从头合成的甲基转移酶建立；②它可以在每次细胞分裂时通过维持甲基转移酶被稳定继承；③它可以在生殖系中被擦除而在下一代中重建，这可由被动去甲基化也可能通过去甲基化酶主动完成。

DNA 甲基化可能潜在两种不同的基因组印记功能。①它可以只存在于一个配子的染色体上来帮助一个亲本等位基因沉默。这是因为 DNA 甲基化与基因抑制相关。为了确定它确切作用，首先必要指明 DNA 甲基化只在一条亲本染色体上存在（也就是一个 DMR）。②有必要确定哪些印记基因以及基因调控装置中哪些部分是被 DNA 甲基化标记的。启动子或长程正负调控元件上甲基化标记的位置对基因表达有不同的作用；最后，有必要明确 DMR 是在哪个发育阶段形成的。如果是在配子发生时期，而且在体细胞的相应位置维持下来（配子 DMR），它可能就是印记。如果它是在胚胎形成二倍体后加到基因上（体细胞 DMR），这时两条亲本染色体在同一个细胞当中了，则不可能是身份标记，却可能对维持亲本特异性沉默有作用。

亲本等位基因特异性 DNA 甲基化在多数印记簇中都有发现。例如 *Igf2* 基因簇有一个位于 *H19* ncRNA 上游 2kb 处的配子 DMR，它只在父源配子中甲基化而其后在所有体细胞组织中都维持着这种状态（Bartolomei et al. 1993；Ferguson-Smith et al. 1993）。类似的配子 DMR 也覆盖了 *Air* ncRNA 的启动子，后者只在沉默的母源基因拷贝中存在，是在雌性配子中获得的（Stoger et al. 1993）。令人惊讶的是，配子 DMR 不是在这些簇（*Igf2*，*Igf2r*）中主要的印记基因启动子区域发现的。相反，沉默的 *Igf2* 启动子没有被甲基化而沉默的 *Igf2r* 启动子处于一个体细胞 DMR 中（Sasaki et al. 1992；Stoger et al. 1993）。相似的，在另外四个深入研究的印记基因簇中（*Pws*、*Kcnq1*、*Gnas* 和 *Dlk1*），发现配子 DMR 在带有印记 ncRNA 的沉默拷贝的染色体上是甲基化的（Shemer et al. 1997；Liu et al. 2000；Takada et al. 2002；Yatsuki et al. 2002）。

体细胞 DMR 的情况不多，只在每个簇中少数蛋白编码的印记基因上出现，暗示 DNA 甲基化可能在维持印记基因表达中的作用有限（Stoger et al. 1993；Moore et al. 1997；Ytsuki et al. 2002）。删除小鼠的配子 DMR 导致多个基因的印记完全丢失，证明这类 DMR 也是整个簇的主要 ICE（图 19-6）（Wutz 1997；Thorvaldsen et al. 1998；Bielinska et al. 2000；Fitzpatrick et al. 2002；Lin et al. 2003；Williamson et al. 2006）。相反，体细胞 DMR 缺失影响了相关印记基因，但印记表达被簇中其他基因维持（Constancia et al. 2000；Sleutels et al. 2003）。

Dnmt 基因家族突变能引起基因组范围的 DNA 甲基化缺失，突显了其调控印记基因表达的重要的作用。从头甲基化酶 Dnmt3a、甲基化酶激活因子 Dmnt3L 或是维持性

甲基化酶 Dnmt1 的突变会产生 DNA 甲基化缺陷型胚胎，它们的印记基因表达发生改变（见第 18 章）。四个印记基因簇（*Igf2*、*Igf2r*、*Kcnq1* 和 *Dlk1*）受此干扰产生的结果，暗示了 DNA 甲基化一般抑制了表达成簇 mRNA 基因的同一亲本染色体上配子 DMR 的作用。因此如果没有 DNA 甲基化，配子 DMR 也无法正常工作，导致了几个印记的蛋白编码基因（包括 *Igf2*、*Igf2r*、*Kcnq* 和 *Dlk1*），在两条亲本染色体上都被抑制了。这意味着这些 mRNA 基因在哺乳动物基因组中的缺省状态是沉默的，需要表观遗传的激活才会表达。值得注意的是，*H19* ncRNA 正常情况下只在带有非甲基化配子 DMR 的染色体上表达，此处却在两条亲本染色体上都有表达（第 18 章）。而在胎盘中显示印记表达的基因有一些例外（Lewis et al. 2004）。

其他表观遗传修饰的类型是否也作为配子印记存在呢？哺乳动物基因组中的表观遗传机制是很多的（第 3 章），所以 DNA 甲基化不可能仅通过印记机制。影响染色质活性状态的组蛋白修饰也可能产生亲本印记，它们也能满足图 19-3 中的多个前提条件。然而至今为止，只有一个 polycomb 复合体的成分蛋白 Eed（催化组蛋白 H3K27 甲基化）影响了一些父源抑制的基因。Eed 对于基因组印记的影响相对于 DNA 甲基化来说小得多，可能只起到维持的作用（Mager et al. 2003）。

DMR 如何被配子甲基化系统所选择呢？将已知的配子 DMR 进行序列比对，并没有发现明显的序列保守性，只有一些含有一系列正向重复序列，这些序列可能有引起 DNA 甲基化的二级结构（Neumann et al. 1995）。DMR 序列相对于其余基因组明显富含 CpG，这很像哺乳动物基因组中超过半数基因启动子含 CpG 岛（第 18 章）。引人注目的是 CpG 岛启动子有一个关键特征是它们常常缺少 DNA 甲基化，而且早期胚胎细胞中有一些的机制来保持启动子 CpG 岛不被甲基化（Antequera 2003）。不过 CpG 岛可以在肿瘤和衰老过程中变成甲基化。

有两个发现阐明了 DNA 甲基化是如何定位到 DMR 序列的。第一个发现是 *H19* 配子 DMR 甲基化的父源特异性是由于母源配子中此区域本来应该进行的甲基化被 CTCF 蛋白阻止（Fedoriw et al. 2004，见 3.6 节）。这可能暗示着 DNA 甲基化系统缺乏特异性。第二个发现是甲基化附属蛋白 Dnmt3L 在雌雄配子中起不同作用。雄配子中，Dnmt3L在逆转座子的甲基化和沉默中起主要作用，而在 DMR 甲基化上作用甚微；相反，在卵子中，它在 DMR 甲基化上起主要作用，而在逆转座子甲基化中没有功能（Bourc'his and Bestor 2006）。逆转座子是可移动的遗传元件，在哺乳动物基因组中大量存在，它可以通过 RNA 中间体复制然后在新的基因组位点插入（Kazazian 2004）。使逆转座子甲基化沉默的蛋白也可以进行 DMR 甲基化，这一点说明基因组防御 DNA 入侵所需的机制也被用于在母源生殖系中产生印记（Barlow1993）。

## 3.6　印记基因簇中两种类型的顺式沉默作用

目前，针对于多个基因簇中控制印记的机制，提出了两类顺式沉默的假说：绝缘子模型（适用于 *Igf2* 簇）以及 ncRNA 介导的沉默模型（适用于 *Igf2r* 和 *Kcnq1* 簇）。虽然没有全部确定，表 19-2 中的大多数基因簇的某些方面都符合之。在 *Igf2* 位点，将位于 *H19* 转录起始点上游 2kb、*Igf2* 转录起点下游 80kb 的配子 DMR 删除，取得突破性

进展，并证明其为绝缘子模型（图 19-7）（Thorvaldsen et al. 1998）。不管删除的区域是从父源或母源遗传来的，删除后，*H19* 和 *Igf2* 均显示出印记的缺失，这就确定了此 DMR 为一个 ICE。后来发现这个 ICE 结合了 CTCF 蛋白（在 β 珠蛋白位点处调节绝缘子活性的蛋白），而其自身则发挥着绝缘子的作用（Bell and Felsenfeld 2000；Hark et al. 2000）。在这种情况下，绝缘子的定义为：处于增强子和启动子之间并阻断二者相互作用的元件。因此，在此位点的印记基因表达模型为：在母源等位基因上，CTCF 结合 ICE 并阻止 *Igf2* 和 *Ins2* 接近与 *H19* ncRNA 共有的增强子（处于三者下游）。这样一来，就只有 *H19* 可以接近这个增强子（图 19-7）。父源等位基因上的 ICE 是在雄性生殖系中得到了 DNA 甲基化的，CTCF 就无法结合上去，父源染色体上的 *Igf2* 和 *Ins2* 与增强子作用，于是可以表达。父源 ICE 的 DNA 甲基化使 *H19* 启动子上又发生了次级甲基化修饰而导致 *H19* 沉默，该机制目前还不知道。在了解了 CTCF 在绝缘子模型中

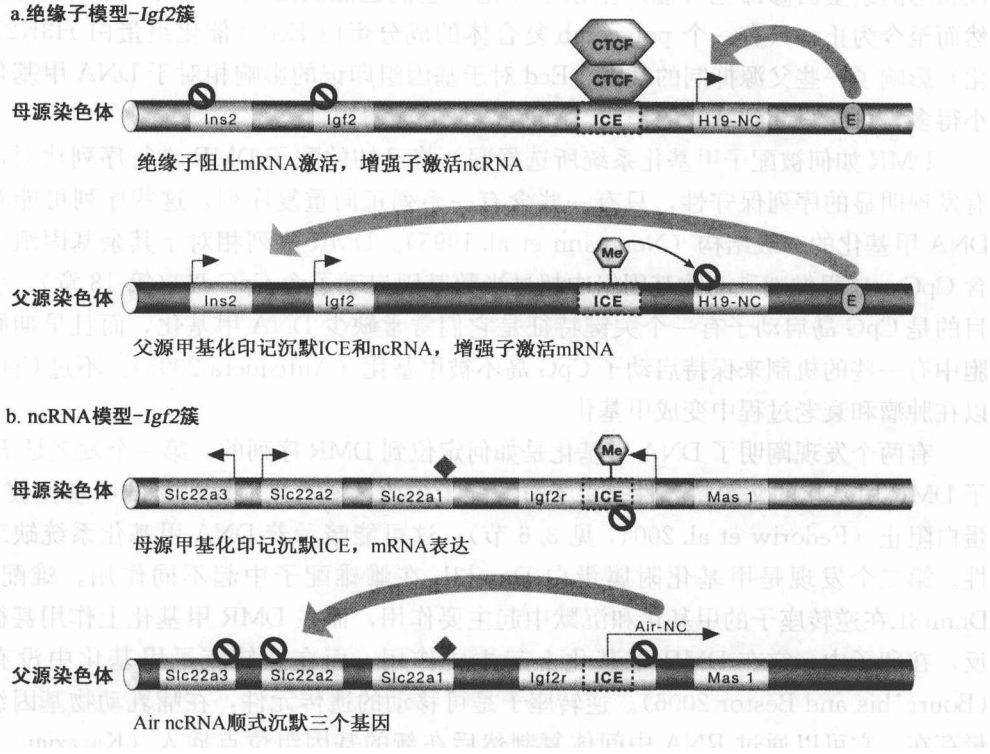



图 19-7　印记基因簇的两个顺式作用沉默机制

(a) *Igf2* 簇的绝缘子模型，显示了内胚层的表达状态。母源染色体上非甲基化的 ICE 结合 CTCF 蛋白形成绝缘子而抑制共同的内胚层增强子（E）并激活 *Igf2* 和 *Ins2*。相反，增强子激活了附近的 *H19* ncRNA 启动子。在父源染色体上甲基化的 ICE 不结合 CTCF，也没有形成绝缘子，所以 *Igf2* 和 *Ins2* mRNA 基因只在此条染色体上表达。*H19* ncRNA 很可能是因为由 2kb 以外的 ICE 传来作用被甲基化，然后被沉默。(b) *Igf2r* 基因簇的 ncRNA 模型。图中所示为胎盘中表达状态。母源染色体上，DNA 甲基化印记直接将其沉默。*Igf2r*、*Slc22a2*、*Slc22a3* mRNA 基因只在这条染色体上表达。*Mas1* 和 *Slc22a1* 在胎盘中不表达。父源染色体上，*Air* ncRNA 启动子处于非甲基化 ICE 中而且表达的，并且顺式沉默了 *Igf2r*，*Slc22a2*、*Slc22a3*。注意在两种模型中，DNA 甲基化印记都使 ncRNA 沉默而 mRNA 得到表达。ICE，印记控制元件；灰色方块，印记 mRNA 基因；基因-NC，印记 ncRNA 基因；箭头，印记基因中的表达等位基因；终止符号，印记基因的抑制等位基因；方块，两条亲本染色体中均沉默的组织特异性基因；箭头，顺式长程效应。

起到的作用之后，人们发现了其他印记基因如 *Rasgrf1*、*Grb10*、*Kcnq1ot1* 上的 CTCF 结合位点，说明绝缘子模型也可能在其他印记基因簇中起作用。

然而，印记模型中的 ncRNA 模式可能更普遍一些。在一个实验中，人们把 108kb 的 *Air* ncRNA 截短到 3kb（Sleutels et al. 2002），从而发现了印记基因簇中的功能性 ncRNA。被截短的 ncRNA 保持着印记的 DNA 甲基化，然而 *Igf2r* 簇中三条 mRNA 基因的沉默都失去了（图 19-7）。现在也发现 ncRNA 调节的沉默机制在 *Kcnq1* 簇中也起作用（Mancini-DiNardo 2006）。此时并不知道 *Air* 和 *Kcnq1ot1* ncRNA 是如何沉默其所在基因簇中的基因的，许多模型都可能参与解释这两个簇。这两个簇中的 mRNA 和 ncRNA 之间有正义-反义的重叠关系，这说明有两种可能性：①mRNA、ncRNA 之间可以形成双链 RNA，引起 RNAi（见第 8 章）。②正义-反义的重叠关系引起两个启动子之间的转录干扰，而这个干扰只影响 mRNA 启动子的转录。两种情况下，第一个事件都是重叠的 mRNA 的转录后沉默，之后受抑制的染色质不断积累，并且蔓延到整个簇，引起转录水平的基因沉默。

当然，也可能是印记 ncRNA 包裹于染色体局部区域，直接募集染色质抑制蛋白到印记簇，作用方式类似于 X 染色体失活中的 *Xist* ncRNA（见第 17 章）。印记 ncRNA、*Xist* ncRNA 调节沉默有着类似之处，最显著的是它们都是顺式作用模式，而且 ncRNA 的表达和多个 mRNA 基因的沉默之间都有正相关；更有人认为二者都是从一个普遍的表观遗传机制进化而来（Reik and Lewis 2005）。如果真是这样，印记 ncRNA 可能通过把被抑制的染色质导向印记基因簇而使基因沉默。不过我们还不清楚印记 ncRNA 怎么行使基因沉默功能，也不知道还有多少其他印记 ncRNA 在基因沉默中起作用。

### 4. 基因组印记——哺乳动物的一种表观遗传调控模型

相对于哺乳动物其他表观遗传调控模式，基因组印记有其优势。有活性和非活性的亲本等位基因都在同一个核中，其转录环境是相同的，所以二者之间任何表观遗传的差异都更有可能是与其转录状态相关联的。“先后”表观遗传系统则相反，它的表观遗传变化也可能反映了细胞分化状态的改变。在有活性的和沉默的亲本等位基因存在于同一个细胞核的情况下，基因组印记就成为研究表观遗传基因调控的理想系统，但是要区分亲本等位基因来辨明基因活性、沉默相关的特异性特征也是个难点。小鼠中能够区别父源和母源染色体的模式系统则在很大程度上克服了这个难题（图 19-1）。

尽管进化过程中表观遗传基因调控途径是高度保守的，但是每个生物体之间也可能有差异，这与其基因组组织不同有关。哺乳动物基因组的组织方式有个特点，就是有高拷贝数重复片段在基因之间起作用，这极大地增加了大多数基因的长度，而且增大了相邻基因间的距离（Kazazian 2004）。其他模式生物则相反，如酵母、原生生物、霉菌、四膜虫、植物和果蝇，它们的基因组都倾向于保持无重复片段的状态，或者至少是在基因中没有。（Rabinowicz et al. 2003）。

在许多生物体中，高拷贝数重复片段可引起 DNA 甲基化、组蛋白修饰抑制，人们认为这是对外源 DNA 序列入侵（如逆转座子、转座子和病毒）的防御性适应。要弄清

表观遗传机制如何在哺乳动物起作用就必须考虑重复片段、基因相互嵌入的本质（Goll and Bestor 2005）。值得注意的是，哺乳动物的内含子富含重复片段，但基因仍然能得到高效表达，所以哺乳动物基因组形成大块沉默异染色质和活性常染色质的可能性很小。人类染色质的广泛研究表明，一般只有启动子或调控元件区域有具活性的组蛋白修饰，这也支持了以上观点（Kim et al. 2005）。印记基因成簇排列，簇中重叠基因相反表达，还有的基因逃避了印记沉默，染色质修饰在哺乳动物基因组中的分布可能相隔不远。

基因组印记对于理解表观遗传学有什么意义呢？印记基因簇还远没有被了解清楚，但显然它们能为了解某一区域或结构域内基因的调控提供信息。现在已经知道有的印记基因簇含有受 DNA 甲基化调控的顺式作用序列；有的所含基因在哺乳动物基因组中被沉默了，需要表观遗传的激活才能表达；有的含有起绝缘子作用的长程调控元件；还有的含有能顺式沉默大片基因结构域的 ncRNA。今后可以证明这些表观遗传调控机制在哺乳动物基因组中是只适用于印记基因簇还是也可以调控非印记基因。

## 5. 展　望

自 1991 年在哺乳动物中发现第一批印记基因以来，基因组印记就成为人们关注的焦点。一些问题仍待解决，尤其是关系到只在哺乳动物中才有的印记基因调控胚胎、子代生长的问题。另一方面，十五年来，对哺乳动物印记表达的表观遗传调控机制的研究取得了重大进展，因此我们认为已经了解了印记机制在印记基因簇的基本作用原理，尽管并不清楚所有的细节问题。

我们知道基因组印记通过细胞的一般表观遗传系统来调控亲本特异表达，而这个系统只在配子中起作用。因此对于一条亲本等位基因来说，所有后续事件都已经被决定了。我们知道不同基因簇的调控机制有着相似之处，但是不知道在哺乳动物基因组中到底有多少这种机制的变种。

印记基因簇受到表观遗传机制的调控，将来我们很想知道非印记基因在多大程度也受到它的控制。我们最终想知道是否可以运用这些知识到人类的医疗事业中。例如，Prader-willi 和 Angelman 综合征患者由于携带印记的等位基因发生缺失会有行为和生长的缺陷，如诱导沉默的等位基因重新表达则能达到治疗目的（Jiang et al. 2004）。了解细胞控制表观遗传信息的途径变得越来越重要，因为在癌症（见第 24 章）、辅助生育技术（见第 22 章）以及衰老（Egger et al. 2004）中也发现有表观遗传调控的紊乱。我们期待对基因组印记的更深入的了解，以及重要模型的提出，以助于解析哺乳动物基因组如何使用表观遗传机制来调控基因表达。

## 致　谢

感谢 Barlow 和 Bartolomei 实验室从前及现在成员对本文提出的观点进行讨论并提出不同意见。由于参考文献数量有限，我们无法提供引用的所有原始数据，在此表示歉意。最后，非常感谢 Marie-Laure Capparros，Anne Ferguson-Smith，Thomas Jenu-

wein，Davor Solter 和 Shirley Tilghman 对初稿提出修订意见。

朱慧君　译
王大亮　校
孙方霖　审校

## 参考文献


Antequera F. 2003. Structure, function and evolution of CpG island promoters. *Cell. Mol. Life Sci.* **60:** 1647–1658.

Barlow D.P. 1993. Methylation and imprinting: From host defense to gene regulation? *Science* **260:** 309–310.

Barlow D.P., Stoger R., Herrmann B.G., Saito K., and Schweifer N. 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature* **349:** 84–87.

Bartolomei M.S., Zemel S., and Tilghman S.M. 1991. Parental imprinting of the mouse H19 gene. *Nature* **351:** 153–155.

Bartolomei M.S., Webber A.L., Brunkow M.E., and Tilghman S.M. 1993. Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. *Genes Dev.* **7:** 1663–1673.

Barton S.C., Surani M.A., and Norris M.L. 1984. Role of paternal and maternal genomes in mouse development. *Nature* **311:** 374–376.

Bell A.C. and Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* **405:** 482–485.

Bielinska B., Blaydes S.M., Buiting K., Yang T., Krajewska-Walasek M., Horsthemke B., and Brannan C.I. 2000. De novo deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. *Nat. Genet.* **25:** 74–78.

Bourc'his D. and Bestor T.H. 2006. Origins of extreme sexual dimorphism in genomic imprinting. *Cytogenet. Genome Res.* **113:** 36–40.

Brannan C.I., Dees E.C., Ingram R.S., and Tilghman S.M. 1990. The product of the H19 gene may function as an RNA. *Mol. Cell Biol.* **10:** 28–36.

Cattanach B.M. and Kirk M. 1985. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315:** 496–498.

Chaillet J.R., Bader D.S., and Leder P. 1995. Regulation of genomic imprinting by gametic and embryonic processes. *Genes Dev.* **9:** 1177–1187.

Chandra H.S. and Nanjundiah V. 1990. The evolution of genomic imprinting. *Dev. Suppl.* **1990:** 47–53.

Constancia M., Dean W., Lopes S., Moore T., Kelsey G., and Reik W. 2000. Deletion of a silencer element in Igf2 results in loss of imprinting independent of H19. *Nat. Genet.* **26:** 203–206.

Cooper D.W., VandeBerg J.L., Sharman G.B., and Poole W.E. 1971. Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. *Nat. New Biol.* **230:** 155–157.

Crouse H.V., Brown A., and Mumford B.C. 1971. Chromosome inheritance and the problem of chromosome "imprinting" in Sciara (Sciaridae, Diptera). *Chromosoma* **34:** 324–398.

Davis E., Caiment F., Tordoir X., Cavaille J., Ferguson-Smith A., Cockett N., Georges M., and Charlier C. 2005. RNAi-mediated allelic *trans*-interaction at the imprinted *Rtl1/Peg11* locus. *Curr. Biol.* **15:** 743–739.

DeChiara T.M., Robertson E.J., and Efstratiadis A. 1991. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* **64:** 849–859.

Egger G., Liang G., Aparicio A., and Jones P.A. 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429:** 457–463.

Fedoriw A.M., Stein P., Svoboda P., Schultz R.M., and Bartolomei M.S. 2004. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science* **303:** 238–240.

Ferguson-Smith A.C., Sasaki H., Cattanach B.M., and Surani M.A. 1993. Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature* **362:** 751–755.

Ferguson-Smith A.C., Cattanach B.M., Barton S.C., Beechey C.V., and Surani M.A. 1991. Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature* **351:** 667–670.

Fitzpatrick G.V., Soloway P.D., and Higgins M.J. 2002. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of *KvDMR1*. *Nat. Genet.* **32:** 426–431.

Goll M.G. and Bestor T.H. 2005. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74:** 481–514.

Hark A.T., Schoenherr C.J., Katz D.J., Ingram R.S., Levorse J.M., and Tilghman S.M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* **405:** 486–489.

Haussecker D. and Proudfoot N.J. 2005. Dicer-dependent turnover of intergenic transcripts from the human β-globin gene cluster. *Mol. Cell Biol.* **25:** 9724–9733.

Hoppe P.C. and Illmensee K. 1982. Full-term development after transplantation of parthenogenetic embryonic nuclei into fertilized mouse eggs. *Proc. Natl. Acad. Sci.* **79:** 1912–1916.

Hurst L.D., McVean G., and Moore T. 1996. Imprinted genes have few and small introns. *Nat. Genet.* **12:** 234–237.

Jiang Y.H., Bressler J., and Beaudet A.L. 2004. Epigenetics and human disease. *Annu. Rev. Genomics Hum. Genet.* **5:** 479–510.

Johnson D.R. 1974. Hairpin-tail: A case of post-reductional gene action in the mouse egg. *Genetics* **76:** 795–805.

Kazazian H.H., Jr. 2004. Mobile elements: Drivers of genome evolution. *Science* **303:** 1626–1632.

Kim T.H., Barrera L.O., Zheng M., Qu C., Singer M.A., Richmond T.A., Wu Y., Green R.D., and Ren B. 2005. A high-resolution map of active promoters in the human genome. *Nature* **436:** 876–880.

Kono T., Obata Y., Wu Q., Niwa K., Ono Y., Yamamoto Y., Park E.S., Seo J.S., and Ogawa H. 2004. Birth of parthenogenetic mice that can develop to adulthood. *Nature* **428:** 860–864.

Landers M., Bancescu D.L., Le Meur E., Rougeulle C., Glatt-Deeley H., Brannan C., Muscatelli F., and Lalande M. 2004. Regulation of the large (~1000 kb) imprinted murine *Ube3a* antisense transcript by alternative exons upstream of *Snurf/Snrpn*. *Nucleic Acids Res.* **32:** 3480–3492.

Lewis A., Mitsuya K., Umlauf D., Smith P., Dean W., Walter J., Higgins M., Feil R., and Reik W. 2004. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat. Genet.* **36:** 1291–1295.

Lin M.S., Zhang A., and Fujimoto A. 1995. Asynchronous DNA replication between 15q11.2q12 homologs: Cytogenetic evidence for maternal imprinting and delayed replication. *Hum. Genet.* **96:** 572–576.

Lin S.P., Youngson N., Takada S., Seitz H., Reik W., Paulsen M., Cavaille J., and Ferguson-Smith A.C. 2003. Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the *Dlk1-Gtl2* imprinted cluster on mouse chromosome 12. *Nat. Genet.* **35:** 97–102.

Liu J., Litman D., Rosenberg M.J., Yu S., Biesecker L.G., and Weinstein L.S. 2000. A *GNAS1* imprinting defect in pseudohypoparathyroidism type IB. *J. Clin. Invest.* **106:** 1167–1174.

Lyle R., Watanabe D., te Vruchte D., Lerchner W., Smrzka O.W., Wutz A., Schageman J., Hahner L., Davies C., and Barlow D.P. 2000. The imprinted antisense RNA at the *Igf2r* locus overlaps but does not imprint *Mas1*. *Nat. Genet.* **25:** 19–21.

Mager J., Montgomery N.D., de Villena F.P., and Magnuson T. 2003. Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat. Genet.* **33:** 502–507.

Mancini-DiNardo D.S., Levorse J,M., Ingram R.S., and Tilghman S.M. 2006. Elongation of the Kcnqot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev.* **20:** 1268–1282.

Mattick J.S. 2005. The functional genomics of noncoding RNA. *Science* **309:** 1527–1528.

McGrath J. and Solter D. 1984a. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37:** 179–183.

———. 1984b. Maternal Thp lethality in the mouse is a nuclear, not cytoplasmic, defect. *Nature* **308:** 550–551.

McLaren A. 1979. *Maternal effects in development: The fourth symposium of the British Society for Developmental Biology* (ed. D.R. Newth and M. Balls). Cambridge University Press, Cambridge, United Kingdom.

Mitsuya K., Meguro M., Lee M.P., Katoh M., Schulz T.C., Kugoh H., Yoshida M.A., Niikawa N., Feinberg A.P., and Oshimura M. 1999. *LIT1*, an imprinted antisense RNA in the human *KvLQT1* locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. *Hum. Mol. Genet.* **8:** 1209–1217.

Moore T. and Haig D. 1991. Genomic imprinting in mammalian development: A parental tug-of-war. *Trends Genet.* **7:** 45–49.

Moore T., Constancia M., Zubair M., Bailleul B., Feil R., Sasaki H., and Reik W. 1997. Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse *Igf2*. *Proc Natl. Acad. Sci.* **94:** 12509–12514.

Neumann B., Kubicka P., and Barlow D.P. 1995. Characteristics of imprinted genes. *Nat. Genet.* **9:** 12–13.

O'Neill M.J. 2005. The influence of non-coding RNAs on allele-specific gene expression in mammals. *Hum. Mol. Genet.* **14:** R113–120.

Rabinowicz P.D., Palmer L.E., May B.P., Hemann M.T., Lowe S.W., McCombie W.R., and Martienssen R.A. 2003. Genes and transposons are differentially methylated in plants, but not in mammals. *Genome Res.* **13:** 2658–2664.

Regha K., Latos P.A., and Spahn L. 2006. The imprinted mouse *Igf2r/Air* cluster—A model maternal imprinting system. *Cytogenet. Genome Res.* **113:** 165–177.

Reik W. 1989. Genomic imprinting and genetic disorders in man. *Trends Genet* **5:** 331–336.

Reik W. and Lewis A. 2005. Co-evolution of X-chromosome inactivation and imprinting in mammals. *Nat. Rev. Genet.* **6:** 403–410.

Reik W., Howlett S.K., and Surani M.A. 1990. Imprinting by DNA methylation: From transgenes to endogenous gene sequences. *Dev. Suppl.* **1990:** 99–106.

Sasaki H., Jones P.A., Chaillet J.R., Ferguson-Smith A.C., Barton S.C., Reik W., and Surani M.A. 1992. Parental imprinting: Potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (Igf2) gene. *Genes Dev.* **6:** 1843–1856.

Schmidt J.V., Levorse J.M., and Tilghman S.M. 1999. Enhancer competition between *H19* and *Igf2* does not mediate their imprinting. *Proc. Natl. Acad. Sci.* **96:** 9733–9738.

Searle A.G. and Beechey C.V. 1978. Complementation studies with mouse translocations. *Cytogenet. Cell Genet.* **20:** 282–303.

Seitz H., Royo H., Lin S.P., Youngson N., Ferguson-Smith A.C., and Cavaille J. 2004. Imprinted small RNA genes. *Biol. Chem.* **385:** 905–911.

Shemer R., Birger Y., Riggs A.D., and Razin A. 1997. Structure of the imprinted mouse *Snrpn* gene and establishment of its parental-specific methylation pattern. *Proc. Natl. Acad. Sci.* **94:** 10267–10272.

Sleutels F., Zwart R., and Barlow D.P. 2002. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415:** 810–813.

Sleutels F., Tjon G., Ludwig T., and Barlow D.P. 2003. Imprinted silencing of *Slc22a2* and *Slc22a3* does not need transcriptional overlap between *Igf2r* and *Air*. *EMBO J.* **22:** 3696–3704.

Spahn L. and Barlow D.P. 2003. An ICE pattern crystallizes. *Nat. Genet.* **35:** 11–12.

Stoger R., Kubicka P., Liu C.G., Kafri T., Razin A., Cedar H., and Barlow D.P. 1993. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell* **73:** 61–71.

Surani M.A., Barton S.C., and Norris M.L. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308:** 548–550.

Takada S., Paulsen M., Tevendale M., Tsai C.E., Kelsey G., Cattanach B.M., and Ferguson-Smith A.C. 2002. Epigenetic analysis of the *Dlk1-Gtl2* imprinted domain on mouse chromosome 12: Implications for imprinting control from comparison with *Igf2-H19*. *Hum. Mol. Genet.* **11:** 77–86.

Thorvaldsen J.L., Duran K.L., and Bartolomei M.S. 1998. Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes Dev.* **12:** 3693–3702.

Tierling S., Dalbert S., Schoppenhorst S., Tsai C.E., Oliger S., Ferguson-Smith A.C., Paulsen M., and Walter J. 2005. High-resolution map and imprinting analysis of the *Gtl2-Dnchc1* domain on mouse chromosome 12. *Genomics.* **87:** 225–235.

Varmuza S. and Mann M. 1994. Genomic imprinting—Defusing the ovarian time bomb. *Trends Genet.* **10:** 118–123.

Verona R.I., Mann M. R., and Bartolomei M.S. 2003. Genomic imprinting: Intricacies of epigenetic regulation in clusters. *Annu. Rev. Cell Dev. Biol.* **19:** 237–259.

Wilkins J.F. and Haig D. 2003. What good is genomic imprinting: The function of parent-specific gene expression. *Nat. Rev. Genet.* **4:** 359–368.

Williamson C.M., Turner M.D., Ball S.T., Nottingham W.T., Glenister P., Fray M., Tymowska-Lalanne Z., Plagge A., Powles-Glover N., Kelsey G., et al. 2006. Identification of an imprinting control region affecting the expression of all transcripts in the *Gnas* cluster. *Nat. Genet.* **38:** 350–355.

Wroe S.F., Kelsey G., Skinner J.A., Bodle D., Ball S.T., Beechey C.V., Peters J., and Williamson C.M. 2000. An imprinted transcript, antisense to *Nesp*, adds complexity to the cluster of imprinted genes at the mouse *Gnas* locus. *Proc. Natl. Acad. Sci.* **97:** 3342–3346.

Wutz A., Smrzka O. W., Schweifer N., Schellander K., Wagner E. F., and Barlow D. P. 1997. Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature* **389:** 745–749.

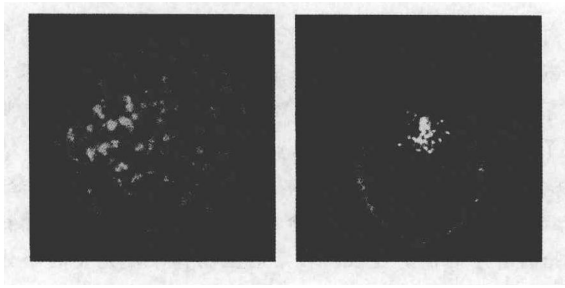
Yatsuki H., Joh K., Higashimoto K., Soejima H., Arai Y., Wang Y., Hatada I., Obata Y., Morisaki H., Zhang Z., et al. 2002. Domain regulation of imprinting cluster in Kip2/Lit1 subdomain on mouse chromosome 7F4/F5: Large-scale DNA methylation analysis reveals that DMR-Lit1 is a putative imprinting control region. *Genome Res.* **12:** 1860–1870.

Zwart R., Sleutels F., Wutz A., Schinkel A. H., and Barlow D. P. 2001. Bidirectional action of the *Igf2r* imprint control element on upstream and downstream imprinted genes. *Genes Dev.* **15:** 2361–2366.

## WWW Resources

Beechey C.V., Cattanach B.M., Blake A., and Peters J. 2005. MRC Mammalian Genetics Unit, Harwell, Oxfordshire. World Wide Web Site—Mouse Imprinting data and references (http://www.mgu.har.mrc.ac.uk/research/imprinting/). http://nocode.bioinfo.org.cn/ NONCODE—the database of noncoding RNA

# 第 20 章



# 生殖细胞系和多能干细胞

M. Azim Surani[1] and Wolf Reik[2]

[1] *WellcomeTrust Cancer Research UR Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, United Kingdom*

[2] *The Babraham Institute, Babraham Research Campus, Cambridge CB2 4AT, United Kingdom*

## 本章目录

1. 前言：哺乳动物的生活——遗传和表观遗传的统一体
2. 调节生殖细胞决定的遗传和表观遗传机制
   - 2.1 不同物种动物的生殖细胞发育原理
   - 2.2 哺乳动物的早期生殖细胞发育
   - 2.3 *Blimp1* 在 PGC 决定过程中的作用
   - 2.4 在生殖细胞中对体细胞程序的抑制——一个进化上保守的现象
   - 2.5 小鼠中 PGC 分化后的表观遗传重编程调节
   - 2.6 生殖系和干细胞——一种可逆的表型
   - 2.7 从多能性 ES 细胞发育成生殖细胞
   - 2.8 从原始生殖细胞到配子
3. 从卵到早期胚胎
   - 3.1 母源性继承和潜在的非对称性分布？
   - 3.2 受精过程中的表观遗传事件
   - 3.3 从受精卵到囊胚
4. 从多能干细胞到体细胞再回到生殖细胞
   - 4.1 多能干细胞的获得
   - 4.2 多能细胞系的表观遗传学特征
   - 4.3 干细胞的重编程能力
5. 展望

致谢

参考文献

## 概要

卵或卵母细胞是体内最不寻常的细胞，因为它是唯一具有分化成整个生物体潜能的细胞。在1651年，Willian Haarvey最早认识到这一点，并提出了“所有生物体都来自于卵”的论断。他认识到受精卵可以逐步发育成完整的生物体，这一理论的提出为后人对“表观遗传”或“渐进式发育”等概念的建立奠定了基础。这个理论的提出最终导致了“预形成论”（preformationist）的消亡。“预形成论”认为生物个体的发育就是生殖细胞中预先存在的小的完整生物体（所谓侏儒体）的简单长大。后来Conrad Waddington将这个概念描述为“胚胎形成”（epigenetic landscape），形象地代表了生物体从卵开始的有序的发育过程。在某些生物体中，完整的个体可以在不需要雄性个体的情况下从一个卵细胞发育而来，这就是所谓的“孤雌生殖”，但是哺乳动物由于“基因组印记”的存在而不能出现孤雌生殖，因为精子对卵子的受精对于发育成成熟个体是必需的。

在大多数生物体中，发育开始于精子和卵子融合成受精卵，这不仅仅产生一个新的个体，而且至少从理论上导致了无止境的生命繁衍。通过这种方式，生殖细胞在所有世代之间提供了一个持续地联系。这个新受精的卵是独一无二的，因为没有任何其他细胞有这种发育成全新的完整生物体的潜力，这种能力被称为全能性（totipotency）。生殖细胞作为一个整体，可以把遗传信息和表观遗传信息传递给后代，同时为了实现这种潜力，也表现出很多与一般细胞不同的特性。卵细胞还拥有一种惊人的潜力，可以将全能性赋予从体细胞例如神经细胞移植而来的细胞核，这一过程被称做克隆或细胞核重编程。

在受精卵发育的过程中，新分裂出的细胞的全能性会不断下降。在哺乳动物中，只有最初几次分裂产生的细胞具有全能性，原则上可以有能力独立发育成生物体。

早期哺乳动物胚胎发育会产生胚泡（见本章标题左图）。胚泡由将发育成胎盘的滋养外胚层细胞和一群内部的细胞构成，后者将发育成胎儿并最终发育成个体。所以这些内部的细胞会分化出成体拥有的所有的两百多种已知体细胞，这种能力被称为细胞的多能性。在特定培养条件下，这些多能性细胞可以被从早期胚胎中“抢救”回来，可以在体外无限培养并保持其分化成各种胚胎细胞和成体细胞的能力，甚至包括精子和卵子。这样的细胞已经从人和小鼠的胚胎中获得，被称做多能**性胚胎干细胞或胚胎干细胞**。当胚胎着床时，会很快失去其产生多能性胚胎干细胞的能力，并开始程序性的胚胎发育。

在胚胎发育过程中，一种最早分化出的细胞类型是**原生殖细胞**，它是精子和卵子的前体，在标题右图中这些细胞被用绿色标出。这保证了那些最终发育成后代的细胞在胚胎最开始发育的时候就建立起来。原生殖细胞是高度分化的细胞，它们最终会发育成成体中成熟的精子或卵子，以重复生命周期，而其余的体细胞最终都会消亡。因此原生殖细胞是非常特殊的细胞。它们也可以被用来获得多能性干细胞，被称为**胚胎生殖细胞**。

干细胞同样存在于成体中。例如，骨髓中的造血干细胞中分化出数以亿计的各种血细胞。同样的，通过各自的干细胞的分化，我们的皮肤细胞和肠道细胞也不断更新。成体干细胞通常只有分化成一些特定组织中细胞的能力，并不像多能胚胎干细胞那样能分化出所有种类的细胞。一个重要的课题是认清多能胚胎干细胞和成体干细胞的相同点和不同点，包括调节它们特性的最基本的表观遗传机理。对生殖细胞和多能性胚胎干细胞独有的表观遗传特性的理解，最终会使我们在医疗特别是生殖医学领域发展出新的概念。

## 1. 前言：哺乳动物的生活——遗传和表观遗传的统一体

遗传信息编码在每个基因组中，建立于受精过程中。在发育过程中，遗传信息是保持不变的，除非突变或是如免疫系统中发生的正常的序列变化。卵或卵母细胞为胚胎的发育提供了三种可以继承的信息（图 20-1a），这仅是一半的遗传信息，另一半则来自于精子。但同时，作为胚胎发育的宿主，母体通过提供 RNA 和蛋白质等物质来主导某些早期发育事件。最后，卵中的表观遗传信息会影响基因组并对发育过程起调节作用。

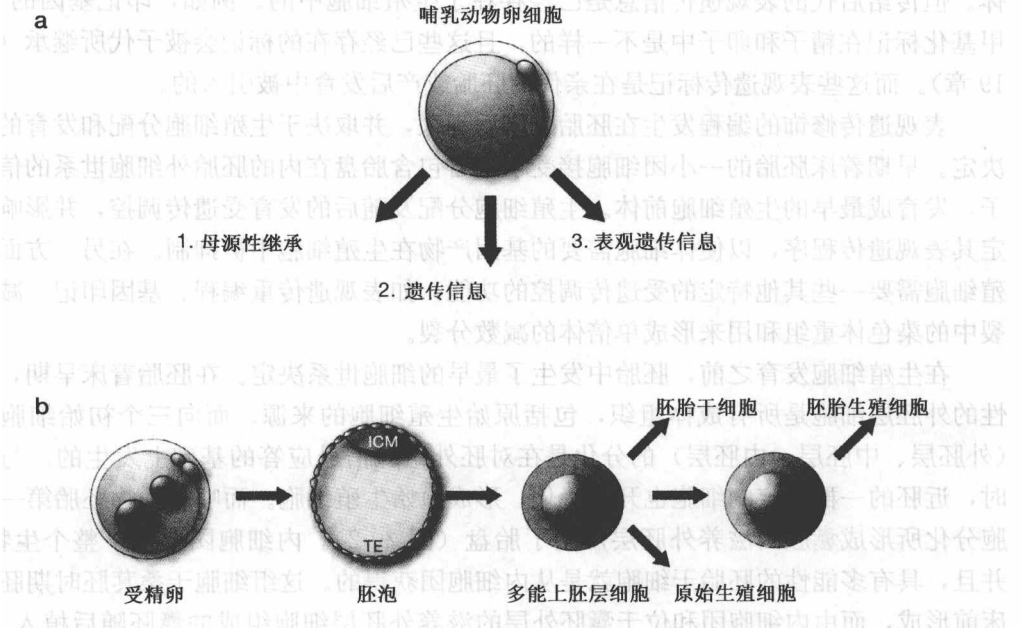



图 20-1　哺乳动物卵母细胞和受精卵

(a) 哺乳动物的卵母细胞包括可以决定早期发育过程的母源 RNA 和母源蛋白质（母源性继承），此外，还有遗传信息（母源染色体）和表观遗传信息（DNA 甲基化和染色体修饰）。(b) 受精卵可以发育成胚泡，胚泡中的内细胞团可以生成用来体外培养的 ES 细胞，也可以发育成所有的体细胞和原始生殖细胞（PGC）。原始生殖细胞可以生成体外培养的 EG 细胞。

表观遗传信息包括 DNA 甲基化、组蛋白修饰、组蛋白变体和非组蛋白类染色质蛋白，它们在发育和分化过程中经历巨大的变化。表观遗传最重要的特征之一就是这些表观遗传标记可以从一代细胞传给下一代细胞，并能够调节基因的转录活性。表观遗传信息因此被认为对建立和维持那些决定细胞命运的基因表达程序的确定性和稳定性有重要意义。在全能性和多能性细胞中，表观遗传标记是不稳定的、可塑的，但在发育的过程中，细胞的这种潜力逐步受到限制，表观遗传标记也越来越稳定并受限制。全能性和多能性细胞，如生殖细胞或胚胎干细胞，都有一个共同的特性，它们可以重编程基因组并抹除表观遗传标记，这种能力可能是其发育可塑性的基础。

发育决定和表观遗传基因调控的互相依赖，在哺乳动物中建立起了遗传和表观遗传

的统一体。这是因为发育的事件，如细胞间的信号传递，会激活一些特殊的基因表达程序，而这种基因表达程序可为表观遗传信息所控制。除此之外，发育过程中还可以建立新的表观遗传事件（如生殖细胞中印记基因的甲基化或去甲基化）。表观遗传标记的建立或抹除反过来可以决定新的基因表达程序，进而影响到每个细胞对发育线索的反应。发育和表观遗传在生殖细胞中的联系更加吸引人，因为这不仅关系到下一代，还很有可能影响到更深远的未来。

生命真的是从受精开始的吗？在遗传上，新的个体确实是从精子和卵子结合的那一瞬间开始的。正是这时，从母本来的单倍体和从父本来的单倍体组成了他们后代的二倍体。但传给后代的表观遗传信息是已经存在于生殖细胞中的。例如，印记基因的 DNA 甲基化标记在精子和卵子中是不一样的，且这些已经存在的标记会被子代所继承（见第 19 章）。而这些表观遗传标记是在亲代的胚胎或产后发育中被引入的。

表观遗传修饰的编程发生在胚胎着床后不久，并取决于生殖细胞分配和发育的遗传决定。早期着床胚胎的一小团细胞接受了来自包含胎盘在内的胚胎外细胞世系的信号分子，发育成最早的生殖细胞前体。生殖细胞分配及随后的发育受遗传调控，并影响和决定其表观遗传程序，以使体细胞需要的基因产物在生殖细胞中被抑制。在另一方面，生殖细胞需要一些其他特定的受遗传调控的功能，如表观遗传重编程、基因印记、减数分裂中的染色体重组和用来形成单倍体的减数分裂。

在生殖细胞发育之前，胚胎中发生了最早的细胞世系决定。在胚胎着床早期，多能性的外胚层细胞是所有成体组织，包括原始生殖细胞的来源。而向三个初始细胞世系（外胚层、中胚层、内胚层）的分化是在对胚外组织信号应答的基础上发生的。与此同时，近胚的一群特定的细胞也开始特化，形成原始生殖细胞。而哺乳动物胚胎第一次细胞分化所形成囊胚的滋养外胚层产生了胎盘（图 20-2）。内细胞团形成了整个生物体，并且，具有多能性的胚胎干细胞就是从内细胞团获得的。这组细胞于桑葚胚时期胚胎着床前形成，而由内细胞团和位于囊胚外层的滋养外胚层细胞组成的囊胚随后植入子宫。在内细胞团和滋养外胚层中发生的表观调控有显著的区别。例如，在胚外组织中，DNA 甲基化的总体水平较低，并且在两种细胞中，印记基因和印记的 X 染色体失活的维持也有所不同。内细胞团和滋养外胚层细胞的发育，与原始生殖细胞相似，都为一个由转录因子和多能性基因参与的遗传程序所决定。这些不同的遗传程序在早期胚胎中最初如何出现，细胞与细胞之间相互作用的意义，以及表观修饰或卵细胞的胞质内因子如何建立这些不同的遗传程序，仍不清楚。

除了生殖细胞将印记基因承载的表观遗传信息携入到胚胎中，在受精过程中，还伴随着其他显著的表观遗传事件的发生。受精卵中精子基因组的甲基化水平先迅速下降，并在随后的细胞分裂过程中又重新获得 DNA 和组蛋白修饰（图 20-2）。受精卵中的母本基因组不发生 DNA 去甲基化，不过随后在早期胚胎细胞分裂过程中，其 DNA 甲基化水平缓慢下降。虽然这两种结果相反且机制不同过程导致胚胎整体上失去配子来源的表观遗传信息，但似乎这些动态的重编程事件与细胞和遗传进程相互作用，从而做出细胞分别走向内细胞团和滋养外胚层世系的最初决定。这样，表观遗传的生命周期无始亦无终，而是始终与遗传程序相互作用，决定内细胞团、滋养外胚层、多能干细胞、生殖

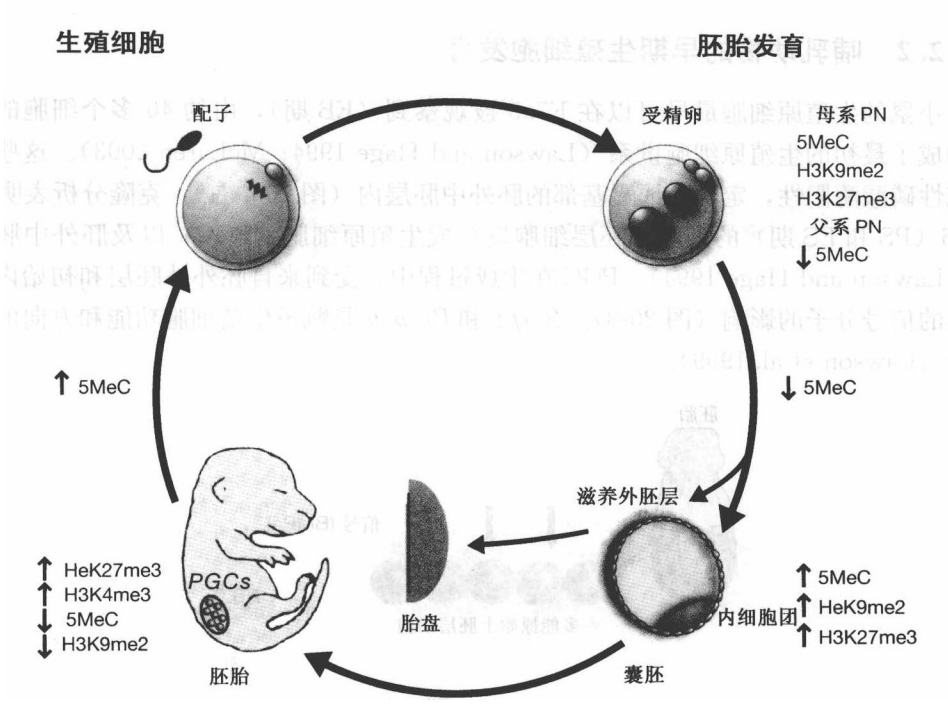



图 20-2　哺乳动物发育过程中的表观遗传重编程循环

受精后，受精卵中的雄原核就包装上组蛋白，但其组蛋白上缺乏 H3K9me2 和 H3K27me3，而此时雌原核则具有上述标记。雄原核基因组会迅速去甲基化，而雌原核基因组则不。被动的去甲基化发生于前着床发育期，直到囊胚期，内细胞团开始出现 DNA 甲基化，H3K9me2 和 H3K27me3 水平的上升。而从滋养外胚层发育而来的胎盘则表现出相对较低的甲基化水平。在进入生殖腺之前和之后，PGC 会逐渐发生 DNA 和 H3K9me2 去甲基化。在生殖细胞发育后期会发生从头的 DNA 甲基化，包括亲本特异的基因印记。

细胞，乃至其他细胞世系的发育过程。

## 2. 调节生殖细胞决定的遗传和表观遗传机制

### 2.1　不同物种动物的生殖细胞发育原理

在动物中，生殖细胞决定是胚胎发育过程中最早的事件之一，使生殖细胞与体细胞得以区分（Surani et al. 2004）。生殖细胞最终会获得细胞全能性。生殖细胞的发育起始有两种重要的模式，“预形成”（这与以前所说的预形成论不是同一概念）和“表观生成”（Extaviyr and Akam 2003）。前者发生于线虫和果蝇中，由某些特殊细胞继承预先形成的生殖细胞决定细胞的特性，产生新的生殖细胞（Leatherman and Jongens 2003；Blackwell 2004）。与之相对应，通过对诱导信号分子的感应，由可能相同的多能细胞中的一部分发育成生殖细胞，另一部分则发育成体细胞，这种机制被称为“表观生成”模式（Lawson and Hage 1994；McLaren 2003）。这一机制控制了小鼠生殖细胞的决定，可能也适用于人类等其他哺乳动物。

## 2.2 哺乳动物的早期生殖细胞发育

小鼠的生殖原细胞最早可以在 E7.5 被观察到（EB 期），由约 40 多个细胞的细胞簇构成了最初的生殖原细胞世系（Lawson and Hage 1994；McLaren 2003）。这些细胞呈碱性磷酸酶阳性，定位于尿囊基部的胚外中胚层内（图 20-4 右）。克隆分析表明E6～E6.5（PS 和 ES 期）的近端外胚层细胞发育成生殖原细胞（PGC）以及胚外中胚层组织（Lawson and Hage 1994）。PGC 在生成过程中，受到来自胚外外胚层和初始内胚层分泌的信号分子的影响（图 20-3）。*Bmp4* 和 *Bmp8b* 是赋予生殖细胞功能和方向的重要分子（Lawson et al. 1999）。

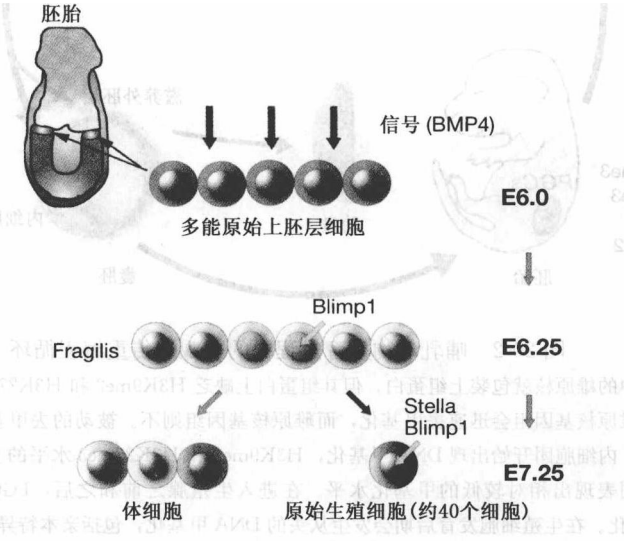



图 20-3 小鼠中的早期生殖细胞决定

从内细胞团（图 20-1 和图 20-2）发育而来的多能性原始上胚层细胞（蓝色）在接受细胞外因子 BMP 刺激后获得产生生殖细胞系（粉色）的能力。活化的 Blimp1 使这些细胞发育成 PGC（红色），而剩下的则发育成体细胞。

为了对 PGC 细胞决定的遗传程序有更细致地了解，人们从初始 PGC 细胞和与之相邻的体细胞中分别制备了单细胞 cDNA（Saitou ta al. 2002）。一系列的标记被用于区分 PGC 细胞和体细胞。通过筛选，两个基因 *fragilis* 和 *stella* 被发现。前者是干扰素诱导的跨膜蛋白家族的新成员，可能与细胞聚集有关，而后者是核质蛋白（既存在于核内也存在于细胞质，译者注）。进一步的研究发现 *fragilis* 表达于 E6.25 的近端外胚层（图 20-3），此时细胞获得了发育成 PGC 和邻近的胚外中胚层细胞的能力。*fragilis* 阳性细胞在原肠胚时期会迁移到近端后部。随即可检测到初始 PGC 细胞，它们表达 *stella*。同时这些最初的 PGC 细胞表达多能性基因，如 *Sox2* 和 *Oct4*，说明这些细胞表现出潜在的多能性，而邻近的体细胞则失去了这种多能性（图 20-3）。相反，初始 PGC 细胞抑制某些基因的表达，包括 *HoxB1* 和 *HoxA1*，而在邻近的体细胞中这些基因在这个时期表达量显著上调。在这些 PGC 细胞中 *Hox* 基因表达的抑制是抑制其走向体细胞命运的重要机制的一部分（Saitou et al. 2002）。

基于对初始 PGC 细胞的分析，很显然，与其他生物体一样，在小鼠 PGC 细胞决定过程中，对体细胞基因程序的抑制是一个重要的特性。（Seydoux and Strome 1999；Blackwell 2004；Surani et al. 2004）。在对基因抑制的候选基因研究中，人们分析组蛋白赖氨酸甲基转移酶（HKMT）以期发现是否某个成员在 PGC 细胞和周围的体细胞之间存在表达差异。在初始 PGC 和体细胞中 *G9a*、*Pfml*、*Setl* 和 *EZH2* 等基因的表达都被检测到。但是，其中的 *Blimpl*（B 细胞成熟诱导蛋白 1）基因在胚胎发育 E7.5 时，只在初始 PGC 中表达，而不在邻近的体细胞中表达。Blimp1 蛋白是一个转录抑制因子，它含有 SET/PR 结构域，一个可以招募 Groucho 和 HDAC 的富集脯氨酸的结构域，可以与 G9a 形成蛋白复合体的 5 个 $C_2H_2$ 锌指结构域，和一个酸性末端（Shaffer et al. 2002；Shapiro-Shelef et al. 2003；Gyori et al. 2004；Sciammas and Davis 2004）。*Blimpl* 最初因其在浆细胞命运决定中的作用被发现，它抑制 B 细胞基因程序，使前体细胞向浆细胞的方向发育（Turner et al. 1994）。事实上，*Blimpl* 在小鼠的发育过程中是广泛表达的（Chang et al. 2002）。

在小鼠早期胚胎中对 *Blimpl* 更详细的分析得到意外的结果。*Blimpl* 的表达开始于胚胎发育 E6.25，也就是原肠胚开始形成时的近外胚层细胞，最初只表达于四到六个和胚外外胚层有直接接触的细胞（图 20-3）（Ohimata et al. 2005）。*Blimpl* 的表达主要集中在短轴的一个末端，这个区域的细胞最终会发育成近侧后部区域。*Blimpl* 蛋白阳性细胞数目逐渐增加，在 mid-streak（MS）期，也就是胚胎发育 E6.75，可以看到在近侧后部区域，20 多个这样的细胞会形成一个紧密地细胞簇。在胚胎发育 E7.5，也就是 EB 期，Blimp1 阳性的细胞会增加到将近 40 个（图 20-4）。这些细胞组成了初始 PGC 细胞，呈经典的碱性磷酸酶阳性，并表达 *Stella*。在遗传世系跟踪实验中确认，开始于胚胎发育 E6.25 的 Blimp1 阳性外胚层细胞确实是世系限制的 PGC 前体细胞。这些数据与先前基于克隆实验的假说相反，先前的实验表明胚胎发育 E6.0～6.5 的细胞并不会世系限制性地发育成 PGC 细胞，因为同一细胞的后代可以发育成体细胞和生殖细胞（Lawson and Hage 1994；McLaren and Lawson 2005）。一个可能的解释是在克隆

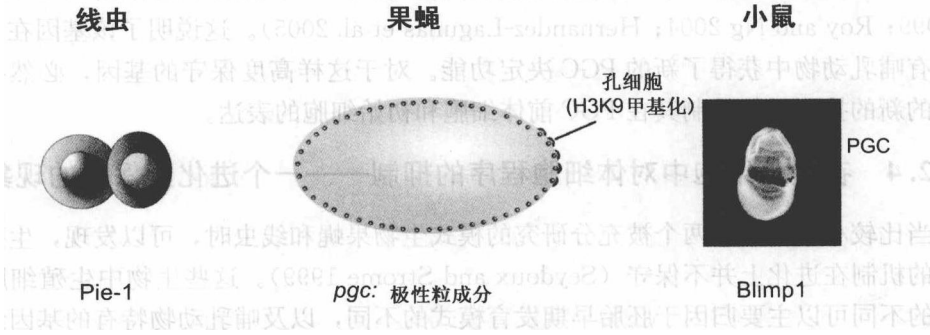



图 20-4　在不同动物中的生殖细胞发育

在线虫中，生殖细胞世系决定发生在受精卵的第一次分裂，生殖细胞系中会表达 *Piel*，使其处于转录静息状态，而另一个细胞则发育成体细胞。在果蝇中，生殖细胞的前体被称作孔细胞，这些孔细胞在受精卵多核体的一侧，其转录静息是依靠 *Pgc* 基因转录出的 RNA 实现的。在小鼠中，最早可见的生殖细胞前体出现在尿囊的基部，这些细胞会表达 *Blimpl*，从而实现转录静息。

分析中，最初标记的细胞是 Blimp1 阴性，然后通过细胞分裂得到一个阳性的细胞，它最终发育成 PGC 细胞，而另一个子细胞则发育成体细胞。调节 Blimp1 阳性细胞数目增加的机制尚不清楚。

## 2.3 *Blimp1* 在 PGC 决定过程中的作用

对 *Blimp1* 基因在 PGC 决定中作用的深入研究对揭开小鼠生殖细胞决定机制产生了深远的影响。*Blimp1* 基因功能缺失表明该基因是小鼠 PGC 决定的重要决定因素（Ohinata et al. 2005；Vincent et al. 2005）。在 E7.5，*Blimp1* 突变体胚胎含有一个异常的由约 20 个 PGC 样的细胞组成的簇，而在正常胚胎中，PGC 继续扩增并向簇外迁移。并且，在 E8.5 突变体胚胎中，这些异常的 PGC 样的细胞的数目没有增加（Ohinata et al. 2005）。

对 *Blimp1* 突变的 PGC 样细胞进行的单细胞分析反映出这些细胞缺乏对 *Hox* 基因的持续抑制。因此，*Blimp1* 很可能起到了抑制初始 PGC 中体细胞程序的作用。在 *Blimp1* 突变体中，PGC 特异基因如 *stella* 和 *Nanos3* 以及一些多能性特异的基因如 *Sox2* 的表达上调也不一致。这些发现强调了 *Blimp1* 在 PGC 决定过程中作为转录调节因子并抑制这些细胞走向体细胞命运的关键作用。

在对 B 细胞的研究中发现 *Blimp1* 通过抑制维持 B 细胞特征性的关键分子，足以诱导细胞分化成为浆细胞（Turner et al. 1994；Shaffer et al. 2002；Shapiro-shelef et al. 2003；Sciammas and Davis 2004；细节详见第 21 章）。Blimp1 通过形成一个 Groucho 和 HDAC2 抑制复合物来发挥作用（Ren et al. 1999）。其锌指结构似乎在与 G9a 形成复合物中也起到了重要作用（Gyory et al. 2004）。但是，在 PGC 决定过程中，由于没有在 *Blimp1* 的 SET/PR 结构域中检测到 HKMT 活性，它是否通过形成一个新的复合物起作用还不清楚。

在脊椎和无脊椎动物中，*Blimp1* 在进化上保守并有许多功能。例如，在脊椎动物如斑马鱼和爪蟾中，*Blimp1* 在数种世系发育，而非 PGC 决定中起作用（de Souza et al. 1999；Roy and Ng 2004；Hernandez-Lagunas et al. 2005）。这说明了该基因在小鼠或者所有哺乳动物中获得了新的 PGC 决定功能。对于这样高度保守的基因，必然有进化而来的新的控制元件控制其在 PGC 前体细胞和初始细胞的表达。

## 2.4 在生殖细胞中对体细胞程序的抑制——一个进化上保守的现象

当比较小鼠与另外两个被充分研究的模式生物果蝇和线虫时，可以发现，生殖细胞决定的机制在进化上并不保守（Seydoux and Strome 1999）。这些生物中生殖细胞决定机制的不同可以主要归因于胚胎早期发育模式的不同，以及哺乳动物特有的基因组印记现象导致的复杂性。但是，重要的是，虽然分子机制有所不同，在生殖细胞决定过程中，对体细胞基因表达程序的抑制这一现象在不同种生物中是共同存在的（Seydoux and Strome 1999；Leatherman and Jongens 2003；Saitou et al. 2003；Blacjwell 2004）。

在线虫中，合子的第一次分裂是不对称的：一个细胞走向体细胞命运（AB），而另一个细胞（P1）走向生殖细胞的命运（图 20-4）。其实，每个 P1、P2、P3 细胞分裂时

都产生一个体细胞，后者进行转录和分化。而将成为生殖细胞的 P1～P3 细胞仍保持转录静息状态。这种转录静息状态是由锌指蛋白 PIE-1 维持的。PIE-1 通过竞争对 RNA 聚合酶Ⅱ CTD 结构域 Ser-2 位点的磷酸化，抑制转录延伸（Van Doren et al. 1998; Seydoux and Strome 1999; Zhang et al. 2003; 详见第 10 章）。但是，体细胞和生殖细胞卵裂球都表现出允许转录的染色质状态，其特征为基因组范围的高水平 H3K4 甲基化。此后，当 P4 卵裂球分裂形成两个生殖系细胞 Z2 和 Z3 时，组蛋白 H3K4 甲基化的减少和抑制性组蛋白 H3K9 甲基化的获得显示抑制型染色质状态开始出现（Scharner et al. 2003）。因此，在线虫生殖细胞系确立的过程中，染色质从转录许可的状态转变为转录沉默的状态。

在果蝇中，生殖细胞系的确立又与小鼠和线虫的情况不同。生殖系的前体细胞叫做孔细胞，在胚胎发育之前，就在受精 syncytial（多核）卵中被检测到了（图 20-4），并且这些细胞像在线虫中观察到的一样，由于 RNA 聚合酶Ⅱ CTD 没有磷酸化，也是转录静息的（Seydoux and Dunn 1997; Van Doren et al. 1998; Schaner et al. 2003）。虽然这种转录沉默也伴随着抑制性染色质修饰，H3K9 甲基化，孔细胞最终只发育成生殖细胞。因此这些细胞与秀丽线虫中的 Z2 和 Z3 细胞等价，它们也只能发育成生殖细胞。此外，孔细胞的转录沉默显然受到 *PGC*（*polar granule component*）基因的调节，因为 *pgc* 基因的缺失导致抑制作用的消失，尽管孔细胞依然能被检测到。这些突变的孔细胞具有 Pol Ⅱ CTD 第二位丝氨酸的磷酸化（Deshpande et al. 2004; Martinho et al. 2004）。这显示 *pgc* 基因可能限制了 Pol Ⅱ CTD 磷酸化必需的重要组分，从而阻断转录从预起始复合体到延伸复合体的转变。

对小鼠、果蝇和线虫发育过程中生殖细胞世系决定的分析，清楚地表明转录抑制对细胞避免向体细胞分化起重要作用，虽然实现这一目的的具体机理在这三种模式动物中差异显著。这显然是由于不同的物种具有不同的早期发育事件。

## 2.5　小鼠中 PGC 分化后的表观遗传重编程调节

在 PGC 细胞决定后，生殖细胞中依然存在着广泛的表观遗传编程和重编程（Seki et al. 2005）。通过这个时期的发育，生殖细胞中的某些抑制性表观遗传修饰被擦除，从而使生殖细胞获得重要的多能性特质，这可能是随后产生全能性的先决条件。

可观察到的重要变化之一是，H3K9me2 在 E8 被抹除，以及 E9 时 HP1 α 水平在常染色质区和着丝粒周边易染色质区的下降（Seki et al. 2005）。同时，自 E8 起，PGC 细胞中 DNA 甲基化水平下降。当 DNA 甲基化和 H3K9 的二甲基化水平下降时，另一转录抑制性表观遗传修饰，由 Polycomb 基因 *Ezh2* 介导的 H3K27me3 水平持续上升（图 20-5）。DNA 甲基化水平下降伴随着从头甲基化酶 *DNMT3a* 和 *DNMT3b* 的转录抑制，以及维持性甲基化酶 *DNMT1* 表达量的瞬时下调。值得注意的是，组蛋白 H3K9 二甲基化和 DNA 甲基化水平的下降与恰巧伴随着多能性相关基因 *Nanog* 的表达（Yamaguchi et al. 2005）。*Nanog* 最初表达于桑椹胚的内细胞和囊胚的内细胞团中。但是 *Nanog* 的表达水平在着床后迅速下降，然后在 PGC 决定完成后又重新表达。与其他的多能性基因 *Oct4*、*Sox2* 和 *Esg1* 一起在生殖细胞中表达，表明生殖细胞获得了细胞多能性的特质。

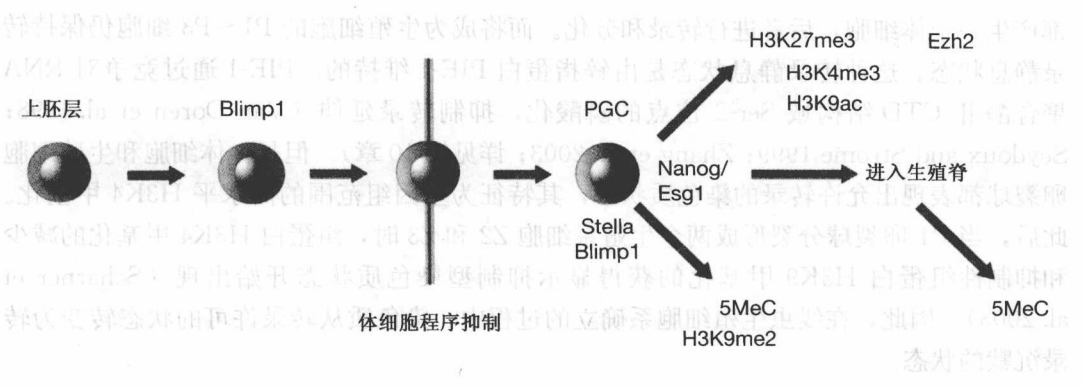



图 20-5 在生殖细胞特化过程中的早期表观遗传事件

囊胚子细胞中的 *Blimp1* 抑制体细胞基因的表达（红色）。随后，*Stella*、*Nanog* 和 *Esg1* 表达；H3K27me3、H3K4me3 和 H3K9ac 上升；H3K9me 和 DNA 甲基化下降。PGC 细胞进入生殖脊，然后发生表观遗传重编程。

当 PGC 细胞进入发育中的生殖脊，更多广泛的表观遗传编程事件继续发生（Surani et al. 2004）。首先，H3K4 甲基化和 H3K9 乙酰化水平上升，这是活跃染色质的特性，以排斥 H3K9me。另外，全基因组范围发生 DNA 去甲基化，抹除亲本印记和单拷贝基因的甲基化。雌性胚胎中，失活的 X 染色体会在这一时期被重新激活。

虽然这样就有一种有效的机制可以抹除“获得的”表观遗传学修饰，但并非所有表观遗传标记都在生殖细胞发育过程中被完全除去。例如，IAP 逆转座子家族的 DNA 甲基化只是部分被消除（Lane et al. 2003）。在配子发生过程中某些表观遗传标记不完全被抹除显然导致生殖系中的表观遗传标记继承，这已在哺乳动物中有许多例证（Chong and Whitelaw 2004）。这种现象究竟有多普遍，又有多少基因座可能与此相关仍需确定。

## 2.6 生殖系和干细胞——一种可逆的表型

生殖细胞和 ICM 细胞是已经证明可以从中获得并培养出多能干细胞的两种细胞。当在有 FGF2 因子的条件下培养时，PGC 可以发生变化从而形成多能性胚胎生殖（EG）细胞（Matsui et al. 1992；Resnick et al. 1992），而 EG 细胞在许多方面与多能性 ES 细胞相类似，除了 EG 细胞在其获得时表现出亲本印记的广泛消除（Tada et al. 1998）。

因为 PGC 细胞表现了一定的多能性特征，可能存在有某种机制使其保有特定的世系特征性。这一点是如何实现的仍不清楚，但是有可能 *Blimp1* 在 PGC 决定后仍然继续起作用，以确保这一进程。而在 EG 细胞获得过程中，可能这一限制被打破，从而使 PGC 获得总体的多能性，能够分化成多种细胞，这在体内的生殖细胞中罕有发生。值得一提的是从 E11.5 到 E12.5 的 PGC 中获得 EG 细胞的效率逐渐降低，进一步的表明在 E11.5 时这些细胞的特征性开始发生改变，从而起始了它们向最终的雄性和雌性生殖细胞的分化之路。

## 2.7 从多能性 ES 细胞发育成生殖细胞

将多能性干细胞引入囊胚，可以使其分化形成所有的体细胞类型，包括生殖细胞

(见下文，图 20-7)。越来越多的努力正试图使 ES 细胞更有效地在体外分化成不同组织。最近，实验证明从培养的 ES 细胞中可以诱导产生 PGC 和类似精子及卵子的结构(Hubner et al. 2003，Toyooka et al. 2003；Geijsen et al. 2004)。因而，随着对 PGC 和配子特化中遗传程序的深入了解，生殖细胞特化的机制可以在体外被验证。这也为研究该世系中的表观遗传学重编程调节过程提供了一个模式系统。这一思路最终有助于我们对人类生殖细胞世系的了解。而且，如果能从培养的 ES 细胞直接诱导分化成人的卵母细胞，就可以利用它们进行“治疗性”克隆，从而绕过对难以获得的供体卵母细胞的需求。这些卵母细胞可以用来进行体细胞核移植而产生囊胚，进而获得 ES 细胞。体细胞核在移植入卵母细胞时发生重编程，从而获得全能性(核移植和重编程详见第 22 章)。这个过程可能对生物医药产生重大影响，因为它可以产生来自于患者本人的 ES 细胞。同时，这个技术使获得大量的，来自不同患者的 ES 细胞库成为可能。这些 ES 细胞可以为研究人类疾病的潜在成因提供机会。例如，它们可以被用于测试各种缓解疾病的治疗性化合物。

人类胚胎和胚胎干细胞在研究和治疗上的应用引发了许多伦理学争议。在不同的国家存在着各种各样的伦理学法规用来监督这些领域的研究。如果能够从 ES 细胞得到活的配子，且其能够受精和发育的话，这必定会引起很多关于如何甚至是否将这项研究应用于人类医学的伦理学问题。

## 2.8　从原始生殖细胞到配子

生殖细胞系发育的下一环节是配子发生的起始和生殖细胞进入减数分裂。生殖嵴中的体细胞环境调节了这一事件的发生时间。在雌性中，生殖细胞停滞于减数分裂前期，而雄性生殖细胞则停滞于有丝分裂。一系列环境信号决定了生殖细胞是否进入减数分裂。最近，一个新基因，*Meisetz*，被证明在减数分裂起始中起关键性作用(Hayashi et al. 2005)。这个基因也含有 SET/PR 结构域和许多锌指结构，被证明具有 H3K4me3 的催化活性。*Meisetz* 在生殖细胞中特异性表达，其表达可在生殖细胞进入减数分裂前期时，即 E13.5 期的雌性和出生后的睾丸中被检测到。*Meisetz* 突变会导致雄性和雌性的不育，表明了其在生殖细胞中的关键作用。突变的生殖细胞表现出显著的 DNA 双链断裂修复途径的缺陷，以及减数分裂中同源染色体配对的缺陷。这项研究表明了生殖细胞减数分裂中表观遗传学机制的重要作用。

在精子发生过程中，大量的表观遗传学修饰继续进行，并且最终将体细胞中的连接组蛋白用睾丸特异性的变体所取代(Kimmins and Sassone-Corsi 2005)，随即，绝大部分组蛋白被精蛋白所取代。研究表明 H3K9 组蛋白甲基转移酶 Suv39 参与了基因表达抑制和染色体配对。并且，两个 SET 结构域蛋白，Suv39h1 和 Suv39h2 在雄性生殖细胞中起作用。后者主要在睾丸中表达，并且富集于性小体的染色质中。Suv39h1 和 Suv39h2 双突变造成精原细胞的发育停滞从而导致不育(Peters et at. 2001)。此外，在雄性生殖细胞中有一个云状结构的染色小体。该结构是一个 RNA 加工小体，由 Dicer 和 Argonaute 蛋白和 micro RNA 组成，是生殖细胞特有的胞质结构。该小体与细胞核相互作用，含有压缩的 mRNA。

生殖系干细胞的自我更新也暗示了非编码 RNA，可能通过 RNA 干涉机制和

HMTase 作用机制，在精子发生中发挥一定的作用。已有报道称 Piwi/Argonaute 家族中的成员（在小鼠中称为 Miwi）在 RNA 干涉中起作用。类 Miwi（Mili）蛋白的丢失导致雄性不育（Luramochi-Miyagawa et al. 2004），并引起逆转座子转录产物 *IAP* 和 *Line1* 的表达升高。但是类 Miwi 蛋白仍未被证明直接参与对上述产物的抑制。

近期的研究成功地从新生乃至成体睾丸中的精原干细胞中获得了多能干细胞（Kanatsu-Shinohara et al. 2004；Guan et al. 2006）。这些细胞可以在体外无限培养，但与 ES 细胞不同的是，它们拥有父本的印记。它们可以在体内以及体外分化成多种类型的体细胞，并且在体内是具活性的生殖细胞。这些细胞可以为研究精子发生的许多方面以及调节其干细胞特征和分化形成精子过程中的表观遗传调节机制提供重要的工具。

在早期生殖细胞中印记的消除产生了哺乳动物生命中第一次也是唯一一次表观遗传学上相等价的亲本染色体。将该种无印记的细胞核直接移植入卵细胞中会因为没有适当的表观遗传学修饰而造成印记基因的错误表达，导致胚胎发育异常并死于早期胚胎发育阶段。这个实验也显示了印记不能由无印记的，直接植入卵细胞的细胞核获得。DNA 甲基化印记的起始开始于出生后雌性生殖细胞的卵发育过程。对于雄性生殖细胞，则发生于胚胎发育的后期。从头甲基转移酶 Dnmt3a 和其辅助因子 Dnmt3L 在这个过程中起到了关键的作用（详见第 19 章）。印记是哺乳动物孤雄生殖的主要障碍。对雌性配子表观基因组操纵的尝试可能会使哺乳动物孤雌胚胎发育成为可能。

## 3. 从卵到早期胚胎

在前面的章节中，我们已经了解到在配子发生过程中成熟的精子和卵子是怎样获得非常特异且不同的表观遗传学标记的。其中的一些差异，如亲本印记，在受精后的胚胎发育过程中始终被忠实的维持。而许多其他标记则在胚胎基因组获得全能性的过程中，发生显著的重编程。从配子中继承的表观遗传学标记在胚胎最早期的分化事件中起着怎样的作用是一个非常重要的问题。如果生命开始于单个具有完整基因组的细胞（即受精卵），其随后进行分裂，那么在子代细胞中基因表达的分化和发育程序究竟是怎样实现的呢？

### 3.1 母源性继承和潜在的非对称性分布？

在具有大的卵细胞的生物中（如果蝇、爪蟾或鸡），一些母源的蛋白或 RNA 在卵中呈不对称分布。因而只被一些子代细胞所继承，并进而发育至特定的命运，而其他未继承这些母源决定性因子的细胞则进行不同的发育（Huynh and St Johnston 2004）。这种策略在相对较大的卵中是可行的（如果蝇），而在相对较小的哺乳动物的卵中比较困难。但是，发育程序并非简单地由卵的大小所决定，更重要的是在哺乳动物中需要生成植入子宫的囊胚和形成维持胚胎发育的胎盘。ICM 在发育上可以被认为与果蝇的卵等价，因为其在早期发育中进行了模式变化以应答于来自胚外体细胞组织的信号。虽然近期研究将受精卵的对称性与囊胚甚至植入后胚胎的对称性相联系（Gardner 1997；Weber et al. 1999），至今尚未在哺乳动物卵中发现非对称性分布的分化决定性因子。而且，哺乳动物胚胎还表现出显著的发育“调节”能力，即当细胞被移去或被干扰时，补偿性

的生长或细胞运动常会发生以维持正常的胚胎发育（Kelly 1977）。尽管如此，早在 4 细胞期，单个细胞（卵裂球）的性质可能已经存在细微的差别，以决定其发育到 ICM 和 TE 世系。（Fujimori et al. 2003；Piotrowska-Nitsche et al. 2005）。

## 3.2　受精过程中的表观遗传事件

在发育和分化的过程中，体细胞世系获得了非常特异和高度特化的 DNA 甲基化和组蛋白修饰模式。当体细胞核被转移进卵时，这些模式显然很难被去除或逆转（见第 22 章）。卵母细胞和精子也具有特化的表观遗传标记，但却能在受精过程中被非常有效地重编程，从而使胚胎的基因组可以行使其新的功能，即获得其全能性（Reik et al. 2001；Surani 2001）。配子的表观遗传学修饰的一些特征和受精后的表观重编程已经被了解（图 20-2）。卵子和精子的基因组都有相当程度的 DNA 甲基化；以一个特殊的序列家族，逆转座子家族为例，池内 A 颗粒（*intracisternal A particles*，*IAP*）在小鼠基因组中约有 1000 个拷贝，在卵子和精子基因组中都被高度甲基化（Lane et al. 2003）。相反，一些特定的序列，尤其是印记基因的差异性甲基化区域（DMR）则只在卵子或精子中被甲基化（见第 19 章）。

卵子基因组也有高度的组蛋白修饰，既有活化形式的（如 H3K9 乙酰化、H3K4 甲基化），也有抑制形式的（如 H3K9 甲基化、H3K27 甲基化）（Morgan et al. 2005）。在受精前，卵子基因组无转录活性，但含有母源性继承而来的，对胚胎前几次分裂，包括那些重编程事件所必需的转录产物和蛋白质（图 20-1a）。精子基因组是高度特化的，其大部分组蛋白在精子发生时已经被替换成高度碱性的精蛋白，以便将 DNA 包装入致密的精子头部（McLay and Clarke 2003）。目前尚不清楚这些残存的组蛋白含有怎样的修饰以及它们存在于基因组的什么位置，鉴于对这些极少量的染色质区域进行研究的困难性。

在受精后的短暂时期内，重编程事件以高度规律的顺序发生于精子的基因组。精蛋白被很快地去掉并被组蛋白所替代。这个事件很可能是 DNA 复制非依赖性的（RI），包括了组蛋白伴侣 HIRA 依赖性的组蛋白变体 H3.3 的置入（van der Heijden et al. 2005；见第 13 章）。同时，在雄原核中发生基因组范围的包括单拷贝和重复序列的 DNA 去甲基化，但 DNA 去甲基化不发生于父源性甲基化的印记基因（Olek and Walter 1997；Mayer et al. 2000；Oswald et al. 2000；Dean et al. 2001；Santos et al. 2002；Lane et al. 2003）。

在 DNA 复制之前，在雄原核中的组蛋白是乙酰化的（H3 和 H4），具有 H3K4 甲基化，并且很快获得 H3K9 单甲基化和 H3K27 单甲基化（Arney et al. 2002；Santos et al. 2002，2005；Lepikhov and Walter 2004）。H3K9 和 H3K27 的二甲基化和三甲基化则只发生于 DNA 复制之后，似乎与核心组蛋白 H3.1 而非 H3.3 的置入相吻合（Santos et al. 2005）。在第一次有丝分裂时，至今为止分析过的绝大多数组蛋白标记开始变得在母本和父本染色体上非常相似，至少从低分辨率的免疫荧光染色结果上看是如此（Santos et al. 2005）。

这些早期重编程过程中的相关酶活性可能都存在于卵子中，或者以蛋白质的形式或者以可以被快速翻译成蛋白质的 RNA 分子的形式。我们已经提到了 HIRA，但是在

DNA 复制后，CAF-1 为复制依赖性（RD）的组蛋白 H3.1 置入过程所必需。Su（var）酶甲基化 H3K9，Ezh2 和其辅助因子 Eed 甲基化 H3K27（Erhardt et al. 2003；Santos et al. 2005）。在父本基因组中发生的显著的 DNA 去甲基化可能是被一个“主动去甲基化”过程所导致，人们为这一过程提出了许多机制，但确定性的酶尚未得到纯化（Morgan et al. 2005）。哺乳动物中一种可能的去甲基化机制是 Aid/Apobec 酶家族，该家族可以脱去 5-甲基胞嘧啶的氨基生成胸腺嘧啶，其产生的 T：G 错配可以被碱基切除修复通路所修复（Morgan et al. 2004）。植物中的去甲基化途径也发生了由 DNA 糖苷酶 Demeter 起始的碱基切除修复（Gehring et al. 2006）。受精时相似的去甲基化酶也可能存在于卵子的细胞质中。这就提出了为何在父本基因组被去甲基化时母本基因组不发生去甲基化的问题。母本染色体或雌原核必然拥有一些特异的保护机制；在雌原核中甲基化 DNA 和 H3K9 二甲基化的共定位说明这种组蛋白修饰可能是保护机制的一种候选方式（Arney et al. 2002；Santos et al. 2002，2005）。

虽然以上证据主要说明在这个过程中组蛋白修饰在整体水平上更多地是被获得而非被丢失，组蛋白精氨酸甲基化仍可能更具动态性。事实上，通过瓜氨酸化反应去除组蛋白精氨酸甲基化的一个候选者 Padi4 就存在于卵子中（Sarmento et al. 2004）。

在受精过程中的快速染色质变化的主要结果，导致在二细胞阶段父本基因组与母本基因组基本相似。但精子基因组因为去甲基化而在表观上发生显著变化。此外，目前分析的结果并没有排除在该阶段建立的组蛋白修饰中存在基因特异性的差异。

### 3.3 从受精卵到囊胚

进一步重编程过程，尤其是基因组范围的 DNA 甲基化模式变化，从二细胞期经过着床前发育的分裂期，一直延续到胚胎发育至囊胚期（Monk et al. 1987；Howlett and Reik 1991；Rougier et al. 1998）。组蛋白修饰的精确动态过程尚未在小鼠中被阐明，但 DNA 甲基化随着每一次核分裂而逐步递减直到 16 细胞桑椹胚阶段。原因是 Dnmt1 在半保留 DNA 复制的过程中维持 CpG 双核苷酸甲基化的甲基化转移酶，被排除在细胞核外（Carlson et al. 1992）。因此，在每次分裂中，基因组中 50%的 DNA 甲基化被丢失。已被报道的唯一例外是印记基因的 DMR。仍不清楚在该过程中是由一种未知的 Dnmt 维持其甲基化，还是由少量的可以进入核中并特异针对于 DMR 的 Dnmt1 来实现这一功能。值得注意的是，在 8 细胞阶段，Dnmt1 蛋白似乎进入核中参与一个复制周期。如果该 Dnmt1 蛋白被去除（通过在卵子中进行基因敲除），DMR 的甲基化水平确实下降 50%，与其仅在一个复制周期中维持甲基化相一致（Howell et al. 2001）。

在 8～16 细胞阶段，桑椹胚外层的细胞变得扁平并成为上皮；这个过程被称为压缩。这是哺乳动物胚胎发育中第一个明显的分化标志。在随后的 2～3 个分裂周期中，桑椹胚开始腔化（形成一个空腔）并且囊胚以其内细胞团（ICM）和外层的滋养外胚层（TE）所被识别。ICM 细胞继续发育形成胚胎和胎儿的所有细胞，而 TE 细胞则形成绝大多数（而非所有）的胎盘的细胞世系（胚外组织）。在此阶段后的短时间内，另一种上皮层细胞在 ICM 的表面形成，这些是初始内胚层细胞，该细胞同样向胎盘和卵黄囊方向发育，而非胚胎方向。这些非常早期的分配事件的一些遗传决定因子已被了解，如

*Oct4*、*Nanog* 和 *Sox2* 对 ICM 细胞的决定和维持起了非常重要的作用。而 *Cdx2* 是 TE 细胞命运的早期维持所必需的（Nichols et al. 1998；Avilion et al. 2003；Chambers et al. 2003；Mitsui et al. 2003；Niwa et al. 2005）。母源蛋白（存在于卵细胞中），或其基因在早期胚胎发育中的表观遗传学调控对早期细胞命运的决定或维持究竟起多大作用，现在仍未知（Dean and Ferguson-Smith 2001）。

但是，主要的表观重编程事件恰恰发生在这个发育阶段。至少通过免疫荧光检测可以判断，ICM 细胞获得了很高水平的 DNA 甲基化（见章名页左图囊胚中红色的内细胞团），而这种 DNA 甲基化很可能是由从头开始的 DNA 甲基转移酶 Dnmt3b 带来（Santos et al. 2002）。这一现象伴随着分别由 G9a 和 Eset 以及 Ezh2 进行的组蛋白 H3K9 和 H3K27 的甲基化（Erhardt et al. 2003）。虽然 DNA 的从头甲基化对 ICM 细胞的初始建立并不关键，由 Ezh2 和 Eset 催化的组蛋白甲基化却非常重要：这两个基因任一被敲除后，内细胞团细胞便不能正确发育（O'Carroll et al. 2001；Dodge et al. 2004）。

与内细胞团中表观遗传修饰的逐渐增多相反，TE 与晚期胎盘的大部分细胞世系一样，仍保持整体的 DNA 低甲基化状态（Chapman et al. 1984；Santos et al. 2002）。胎盘细胞被认为需要较低的表观遗传稳定性，因为同可以发育成为成体生物的胚胎相比，它们的生命周期毕竟更短。

在这些大规模的基因组水平的表观遗传事件的基础上，更多的位点特异性的重编程事件也发生于这一阶段。在 XX 的雌性胚胎中，父源 X 染色体在卵裂阶段失活，并且在胚外组织（如 TE 和胎盘）中一直维持该失活状态（Huynh 和 Lee 2003；Okamoto et al. 2005）。而在内细胞团中，失活的 X 染色体被重新活化，并在接下来的分化过程中随机失活一条 X 染色体（Mak et al. 2004；详见图 17-4）。从机理上，植入前胚胎中印记的 X 染色体失活是由于父源 X 染色体上非编码 RNA *Xist* 的表达，该 RNA 被认为“包被”了染色体从而导致基因沉默，并确立了抑制性的表观遗传修饰（Heard 2004）。在新形成的 ICM 细胞中，*Xist* 的转录被下调，抑制性的表观遗传学修饰随后丢失，染色体被活化（Mak et al. 2004；Okamoto et al. 2004）。随即，表胚层细胞中发生 X 染色体的随机失活。我们在下一节中将会介绍 ES 细胞会被“冰冻”在 X 染色体重新激活后的时期，因此雌性 ES 细胞系含有两条活化的 X 染色体。

## 4. 从多能干细胞到体细胞再回到生殖细胞

### 4.1　多能干细胞的获得

在前一节里，我们了解到在受精卵，分裂期胚胎和囊胚阶段发生了显著的表观遗传学重编程事件，从而造成内细胞团和滋养外胚层具有不同的表观遗传学模式。我们现在来探讨从囊胚和不同胚层中获得并培养的早期干细胞的遗传学和表观遗传学特征。包括胚胎干细胞（Smith 2001），滋养外胚层干细胞（TS）（Rossant 2001），内胚层干细胞（XEN）（Kunath et al. 2005）和胚胎生殖干细胞（EG）（Matsui et al. 1992）。

这些细胞共同特点是它们可以从完整的胚胎中被分离并在特定的培养条件中建系。

一旦建系，这些细胞可以一直培养且并不表现出衰老的特征。这些细胞也可以在培养过程中进行遗传学改造，并重新被引入活的胚胎中来参与适当的细胞世系的发育过程。

20 世纪 80 年代，哺乳动物胚胎学的最重要发现之一是发展出从内细胞团中获得多能胚胎干细胞的方法。从小鼠囊胚中分离的胚胎干细胞，可以维持传代，并且当将其移植回囊胚时，可以分化成为所有种类的细胞，从而形成嵌合体（图 20-6；Evans and Kaufman 1981；Martin 1981）。尤为显著的是 ES 细胞的后代可以分化形成生殖细胞并产生正常的，完全从 ES 细胞基因型获得的后代。这些，与通过同源重组技术从基因水平上操纵 ES 细胞基因组的能力一起，使得小鼠遗传学有了革命性的进展，并使小鼠成为哺乳动物的遗传模式生物。

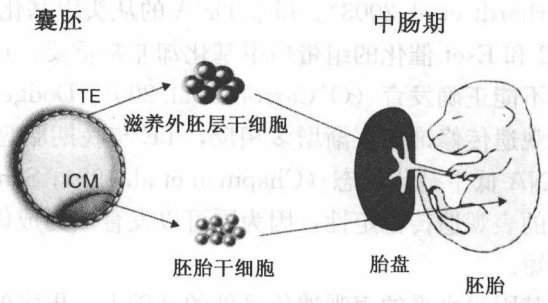



图 20-6 从囊胚中获得胚胎干细胞和滋养外胚层干细胞

从内细胞团中可以分离得到胚胎干细胞，并在体外培养维持不分化状态。这些细胞可以在体外进行基因操作。ES 可以重新被导入囊胚中克隆出胚胎中的任何组织，包括生殖细胞系，但不包括胎盘的滋养外胚层细胞。滋养外胚层干细胞可以通过类似的方法得到，在重新导入囊胚后可以分化成为胎盘细胞。

ES 细胞具有和内细胞团或外胚层细胞相似的特征，但仍与上述两种细胞有相当大的差异，从而使得 ES 细胞成为一种类似于“合成的”而在正常胚胎中并不存在的细胞类型（似乎其他多能性细胞系也是如此）（Smith 2001）。例如，小鼠 ES 细胞的自我更新需要 *Lif/gp130/Stat3* 信号通路，而该信号通路突变的胚胎仍可以发育形成正常的内细胞团（Smith 2001）。因此可能发生了对建立并维持从内细胞团获得的 ES 细胞所必须的表观遗传学变化。内细胞团细胞体外培养时很快就失去 *Oct4* 的表达，只有一种相对容易获得 ES 细胞的小鼠株系，称为 129Sv，可以在体外培养中保留一些表达 *Oct4* 的细胞（Buehr et al. 2003）。在小鼠和恒河猴的 ES 细胞中，也有印记基因的表观遗传学变化的报道，并且在小鼠中，这种变化在重新导入内细胞团所获得的嵌合体中引起细胞的异常发育（Dean et al. 1998；Humpherys et al. 2001；Fujimoto et al. 2005）。

## 4.2 多能细胞系的表观遗传学特征

ES 细胞可以在体外分化成一些不同世系的细胞（图 20-7）。表观遗传学机制在何种程度上维持细胞于未分化或分化的状态？明显的是，未分化的 ES 细胞，分化的 ES 细胞和体细胞有表观基因组的差异，而多能性细胞的特性是其染色质具有高度动态的可塑性（Meshorer et al. 2006）。在 ES 细胞中删除参与控制组蛋白标记变化的 Parp（多聚-ADP 核糖酶），会导致向滋养外胚层细胞进行低水平的转分化事件，这表明表观遗传标

记为维持 ES 细胞的特征性所必需（Hemberger et al. 2003）。DNA 甲基化可能也参与这一事件，因为严重缺失 DNA 甲基化会造成 ES 细胞分化能力的部分障碍，并且使那些能分化的 ES 细胞表达出滋养外胚层的标志基因（Jackson et al. 2004）。

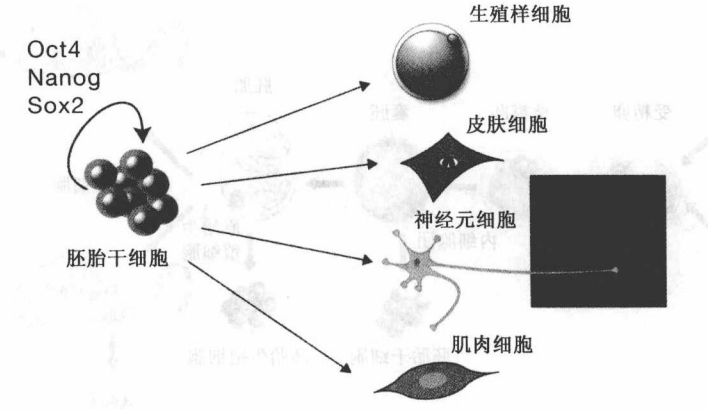



图 20-7　体外分化 ES 到各种细胞类型

ES 可以在体外通过合适的细胞培养条件分化成各种细胞类型，如神经元细胞、肌肉细胞、甚至生殖细胞（卵细胞）。本图中，ES 细胞在体外分化成为神经胶质细胞。

ES 细胞多能性的维持依赖于转录因子 *Oct4*、*Nanog* 和 *Sox2*。这些转录因子单独或共同结合于 ES 细胞中的许多基因位点，而这些基因或表达或被抑制以使多能性得以维持（Boyer et al. 2005；Loh et al. 2006）。ES 细胞体外分化的特征为多能性基因的转录沉默，并在所有体细胞中维持抑制状态。表观遗传机制对多能性基因的抑制的确非常重要：例如，在分化过程中，*Oct4* 的启动子积累抑制性的组蛋白标记和 DNA 甲基化（Feldmann et al. 2006）。再例如，在 *Dnmt1* 敲除的胚胎中，DNA 甲基化的丢失会导致 Oct4 在分化过程中的重新表达（Feldmann et al. 2006）。

ES 细胞及其分化产物也被用来作为 X 染色体失活的表观遗传调节模型。雌性的 ICM 和 ES 细胞有表达下调的 *Xist* 基因和两条活化的 X 染色体。在体外分化过程中，一条 X 染色体的 *Xist* 表达逐渐上调，*Xist* RNA 开始顺式“包被”这条染色体，并且在沉默该 X 染色体上基因的同时聚集抑制性的组蛋白修饰和接踵而来的 DNA 甲基化（Heard 2004）。

其他的多能性细胞类型也可以体外培养建系，但是它们的表观遗传特征不如 ES 细胞明显。然而，在雌性 TS 细胞中对 X 染色体失活的表观遗传研究显示父源 X 染色体失活并含有抑制性组蛋白标记（Huynh 和 Lee，2003）。雌性的 XEN 细胞也有一条父源 X 染色体被失活（Kunath et al. 2005）。

原始生殖细胞（PGC）也可以在其在胚胎中迁移（E8-E10. 5）或在其到达胚胎生殖脊（E11. 5～E13. 5）的时期经体外培养建成多能性细胞系（图 20-8）（Matsui et al. 1992）。在 E8. 5～12. 5 发育时期，PGC 经历了包括印记基因及基因组其他序列 DNA 去甲基化在内的广泛的表观遗传重编程过程（Hajkova et al. 2002；Lee et al. 2002）。大多数 EG 细胞确实经历了印记基因和其他序列的 DNA 去甲基化过程，并且当移植入嵌合体时，因此改变了它

们的发育潜能（Tada et al. 1998）。现在还不清楚的是内源的 PGC 细胞和体外培养的 EG 细胞之间是否像 ICM 和 ES 细胞间一样存在表观遗传差异。

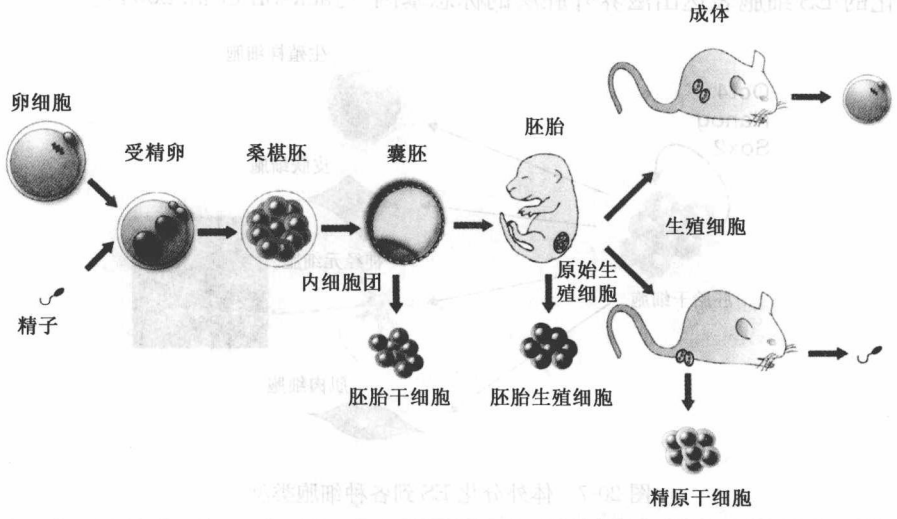



图 20-8 生殖系发育和多能性的联系

卵细胞和受精卵是全能性的（卵因亲本印记基因有一定限制），特定时期内的桑椹胚也是全能性的。多能性细胞可以从内细胞团，PGC 和初生或成体睾丸中的精原干细胞中获得。因此生殖系发育和细胞多能性之间存在着紧密联系。

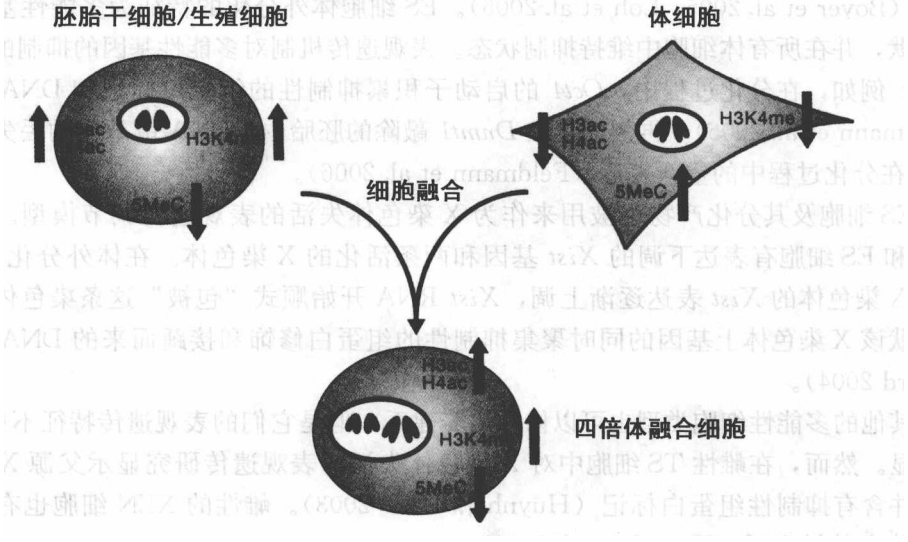



图 20-9 多能性细胞有能力重编程体细胞

ES 或 EG 细胞可以与体细胞融合成四倍体细胞。这会导致体细胞核的表观遗传重编程，例如 DNA 甲基化，H3 和 H4 的乙酰化，H3K4 的甲基化。这些重编程的四倍体融合细胞有一些多能性的表型，他们在注射入囊胚后可以发育成很多细胞类型。

### 4.3　干细胞的重编程能力

干细胞在培养过程中持续保持细胞多能性而不衰老的能力，可能是其表观遗传重编程的必要条件。在细胞融合实验中，这种重编程能力已经被在与体细胞融合的 EG 和 ES 细胞中得到证实（Tada et al. 1997 2001；Cowan et al. 2005）。在这样融合的四倍体细胞中，体细胞的表观遗传模式发生重编程（图 20-9）。在 EG-体细胞的融合细胞中，体细胞的基因组在印记基因和一些其他的基因上发生了 DNA 去甲基化现象（Tada et al. 1997）。在 ES-体细胞的融合细胞中，印记基因的 DNA 没有变化，相反，失活的 X 染色体重新被激活，OCT4 的启动子 DNA 被去甲基化，从而导致 OCT4 的重新表达（Tada et al. 2001；Cowan et al. 2005；Surani 2005）。

到目前为止，对于在这些杂合细胞中是否存在会影响 DNA 甲基化和组蛋白修饰的重编程细胞因子还不清楚。这些 EG 或是 ES 的融合细胞中的去甲基化机理尚不清楚，也许是由于 DNA 去甲基化酶的活性导致的主动去甲基化，也可能是由于复制偶联的 DNMT1 的缺乏导致的被动去甲基化。但无论如何，ES 和 EG 细胞是纯化和鉴定表观遗传重编程细胞因子的重要来源。

## 5. 展　望

在未来的几年中，我们将会看到令人激动的决定性进展，我们会对生殖细胞和干细胞多能性和全能性方面发挥重要作用的遗传和表观遗传因子有更深入的了解。高通量高灵敏度的实验技术将会被用于表观遗传信息的各个层面。能够调节表观遗传信息，特别是能够重编程体细胞表观遗传信息，并促使其向多能性细胞转变的细胞因子将被发现。可以在体内选择性的安全的操作表观遗传状态的新方法也将被开发。

多能干细胞为基础研究提供了令人激动的机会，同时它也是潜在的生物医学的应用技术。在基础研究领域，多能性状态的独特性也许可能揭示对细胞命运决定的机理。利用其能够发育成各种细胞的性质，我们或许可以通过置换细胞的方法修复病原组织。同样地，我们可以通过建立疾病模型，研究细胞是如何通过发育早期的突变和表观突变（如 DNA 甲基化状态或染色体模板改变引起的突变，但不包括 DNA 序列的改变）产生各种疾病。这反过来也许可以用来设计药物甚至预防疾病。

人类多能性干细胞的应用伴随着敏感的伦理问题，这已经引起了世界范围的讨论。建立应用于干细胞医学应用和生物研究的伦理学规则已迫在眉睫，因为在得到多能性干细胞的过程中涉及早期胚胎的操作。

随着对细胞多能性和基因组重编程表观遗传机理的了解，也许将来的某一天我们可以用成体干细胞甚至分化后的细胞重编程得到多能性干细胞，而不用再借助于胚胎操作。也许我们今天对表观遗传机理的理解会对未来医学的进步起到决定性作用。

## 致 谢

M. A. S. 感谢所有现在和过去的实验室人员的原创性工作。这些工作是在 Wellcome Trust、BBSRC、EU、DTI、CellCentric 的资助下完成的。W. R. 感谢实验室中的同事们，特别感谢 Wendy Dean，感谢他们在完成这个章节中的工作和宝贵意见。W. R. 的工作得到了 BBSRC、MRC、EU、DTI 和 CellCentric 的资助。

杨 鹏 王译萱 译
马 奔 校
朱 冰 审校

## 参考文献

Arney K.L., Bao S., Bannister A.J., Kouzarides T., and Surani M.A. 2002. Histone methylation defines epigenetic asymmetry in the mouse zygote. *Int. J. Dev. Biol.* **46:** 317–320.

Avilion A.A., Nicolis S.K., Pevny L.H., Perez L., Vivian N., and Lovell-Badge R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17:** 126–140.

Blackwell T.K. 2004. Germ cells: Finding programs of mass repression. *Curr. Biol.* **14:** R229–230.

Boyer L.A., Lee T.I., Cole M.F., Johnstone S.E., Levine S.S., Zucker J.P., Guenther M.G., Kumar R.M., Murray H.L., Jenner R.G., et al. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122:** 947–956.

Buehr M., Nichols J., Stenhouse F., Mountford P., Greenhalgh C.J., Kantachuvesiri S., Brooker G., Mullins J., and Smith A.G. 2003. Rapid loss of Oct-4 and pluripotency in cultured rodent blastocysts and derivative cell lines. *Biol. Reprod.* **68:** 222–229.

Carlson L.L., Page A.W., and Bestor T.H. 1992. Properties and localization of DNA methyltransferase in preimplantation mouse embryos: Implications for genomic imprinting. *Genes Dev.* **6:** 2536–2541.

Chambers I., Colby D., Robertson M., Nichols J., Lee S., Tweedie S., and Smith A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113:** 643–655.

Chang D.H., Cattoretti G., and Calame K.L. 2002. The dynamic expression pattern of B lymphocyte induced maturation protein-1 (Blimp-1) during mouse embryonic development. *Mech. Dev.* **117:** 305–309.

Chapman V., Forrester L., Sanford J., Hastie N., and Rossant J. 1984. Cell lineage-specific undermethylation of mouse repetitive DNA. *Nature* **307:** 284–286.

Chong S. and Whitelaw E. 2004. Epigenetic germline inheritance. *Curr. Opin. Genet. Dev.* **14:** 692–696.

Cowan C.A., Atienza J., Melton D.A., and Eggan K. 2005. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* **309:** 1369–1373.

Dean W. and Ferguson-Smith A. 2001. Genomic imprinting: Mother maintains methylation marks. *Curr. Biol.* **11:** R527–530.

Dean W., Santos F., Stojkovic M., Zakhartchenko V., Walter J., Wolf E., and Reik W. 2001. Conservation of methylation reprogramming in mammalian development: Aberrant reprogramming in cloned embryos. *Proc. Natl. Acad. Sci.* **98:** 13734–13738.

Dean W., Bowden L., Aitchison A., Klose J., Moore T., Meneses J.J., Reik W., and Feil R. 1998. Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: Association with aberrant phenotypes. *Development* **125:** 2273–2282.

Deshpande G., Calhoun G., and Schedl P. 2004. Overlapping mechanisms function to establish transcriptional quiescence in the embryonic *Drosophila* germline. *Development* **131:** 1247–1257.

de Souza F.S., Gawantka V., Gomez A.P., Delius H., Ang S.L., and Niehrs C. 1999. The zinc finger gene Xblimp1 controls anterior endomesodermal cell fate in Spemann's organizer. *EMBO J.* **18:** 6062–6072.

Dodge J.E., Kang Y.K., Beppu H., Lei H., and Li E. 2004. Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.* **24:** 2478–2486.

Erhardt S., Su I.H., Schneider R., Barton S., Bannister A.J., Perez-Burgos L., Jenuwein T., Kouzarides T., Tarakhovsky A., and Surani M.A. 2003. Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. *Development* **130:** 4235–4248.

Evans M.J. and Kaufman M.H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292:** 154–156.

Extavour C.G. and Akam M. 2003. Mechanisms of germ cell specification across the metazoans: Epigenesis and preformation. *Development* **130:** 5869–5884.

Feldman N., Gerson A., Fang J., Li E., Zhang Y., Shinkai Y., Cedar H., and Bergman Y. 2006. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell. Biol.* **8:** 188–194.

Fujimori T., Kurotaki Y., Miyazaki J., and Nabeshima Y. 2003. Analysis of cell lineage in two- and four-cell mouse embryos. *Development* **130:** 5113–5122.

Fujimoto A., Mitalipov S.M., Kuo H.C., and Wolf D.P. 2005. Aberrant genomic imprinting in rhesus monkey ES cells. *Stem Cells* **24:** 595–603.

Gardner R.L. 1985. Clonal analysis of early mammalian development. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **312:** 163–178.

———. 1997. The early blastocyst is bilaterally symmetrical and its axis of symmetry is aligned with the animal-vegetal axis of the zygote in the mouse. *Development* **124:** 289–301.

Gehring M., Huh J.H., Hsieh T.F., Penterman J., Choi Y., Harada J.J., Goldberg R.B., and Fischer R.L. 2006. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124:** 495–506.

Geijsen N., Horoschak M., Kim K., Gribnau J., Eggan K., and Daley G.Q. 2004. Derivation of embryonic germ cells and male gametes from embryonic stem cells. *Nature* **427:** 148–154.

Guan K., Nayernia K., Maier L.S., Wagner S., Dressel R., Lee J.H., Nolte J., Wolf F., Li M., Engel W., and Hasenfuss G. 2006. Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* **440:** 1199–1203.

Gyory I., Wu J., Fejer G., Seto E., and Wright K.L. 2004. PRDI-BF1 recruits the histone H3 methyltransferase G9a in transcriptional silencing. *Nat. Immunol.* **5:** 299–308.

Hajkova P., Erhardt S., Lane N., Haaf T., El-Maarri O., Reik W., Walter

J., and Surani M.A. 2002. Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117:** 15–23.

Hayashi K., Yoshida K., and Matsui Y. 2005. A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* **438:** 374–378.

Heard E. 2004. Recent advances in X-chromosome inactivation. *Curr. Opin. Cell Biol.* **16:** 247–255.

Hemberger M., Nozaki T., Winterhager E., Yamamoto H., Nakagama H., Kamada N., Suzuki H., Ohta T., Ohki M., Masutani M., and Cross J.C. 2003. Parp1-deficiency induces differentiation of ES cells into trophoblast derivatives. *Dev. Biol.* **257:** 371–381.

Hernandez-Lagunas L., Choi I.F., Kaji T., Simpson P., Hershey C., Zhou Y., Zon L., Mercola M., and Artinger K.B. 2005. Zebrafish *narrow-minded* disrupts the transcription factor *prdm1* and is required for neural crest and sensory neuron specification. *Dev. Biol.* **278:** 347–357.

Howell C.Y., Bestor T.H., Ding F., Latham K.E., Mertineit C., Trasler J.M., and Chaillet J.R. 2001. Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104:** 829–838.

Howlett S.K. and Reik W. 1991. Methylation levels of maternal and paternal genomes during preimplantation development. *Development* **113:** 119–127.

Hubner K., Fuhrmann G., Christenson L.K., Kehler J., Reinbold R., De La Fuente R., Wood J., Strauss J.F., III, Boiani M., and Scholer H.R. 2003. Derivation of oocytes from mouse embryonic stem cells. *Science* **300:** 1251–1256.

Humpherys D., Eggan K., Akutsu H., Hochedlinger K., Rideout W.M., III, Biniszkiewicz D., Yanagimachi R., and Jaenisch R. 2001. Epigenetic instability in ES cells and cloned mice. *Science* **293:** 95–97.

Huynh J.R. and St Johnston D. 2004. The origin of asymmetry: Early polarisation of the *Drosophila* germline cyst and oocyte. *Curr. Biol.* **14:** R438–449.

Huynh K.D. and Lee J.T. 2003. Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. *Nature* **426:** 857–862.

Jackson M., Krassowska A., Gilbert N., Chevassut T., Forrester L., Ansell J., and Ramsahoye B. 2004. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol. Cell. Biol.* **24:** 8862–8871.

Kanatsu-Shinohara M., Inoue K., Lee J., Yoshimoto M., Ogonuki N., Miki H., Baba S., Kato T., Kazuki Y., Toyokuni S., et al. 2004. Generation of pluripotent stem cells from neonatal mouse testis. *Cell* **119:** 1001–1012.

Kelly S.J. 1977. Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. *J. Exp. Zool.* **200:** 365–376.

Kimmins S. and Sassone-Corsi P. 2005. Chromatin remodelling and epigenetic features of germ cells. *Nature* **434:** 583–589.

Kunath T., Arnaud D., Uy G.D., Okamoto I., Chureau C., Yamanaka Y., Heard E., Gardner R.L., Avner P., and Rossant J. 2005. Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132:** 1649–1661.

Kuramochi-Miyagawa S., Kimura T., Ijiri T.W., Isobe T., Asada N., Fujita Y., Ikawa M., Iwai N., Okabe M., Deng W., et al. 2004. *Mili*, a mammalian member of *piwi* family gene, is essential for spermatogenesis. *Development* **131:** 839–849.

Lane N., Dean W., Erhardt S., Hajkova P., Surani A., Walter J., and Reik W. 2003. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35:** 88–93.

Lawson K.A. and Hage W.J. 1994. Clonal analysis of the origin of primordial germ cells in the mouse. *CIBA Found. Symp.* **182:** 68–84.

Lawson K.A., Dunn N.R., Roelen B.A., Zeinstra L.M., Davis A.M., Wright C.V., Korving J.P., and Hogan B.L. 1999. *Bmp4* is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **13:** 424–436.

Leatherman J.L. and Jongens T.A. 2003. Transcriptional silencing and translational control: Key features of early germline development. *Bioessays* **25:** 326–335.

Lee J., Inoue K., Ono R., Ogonuki N., Kohda T., Kaneko-Ishino T., Ogura A., and Ishino F. 2002. Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* **129:** 1807–1817.

Lepikhov K. and Walter J. 2004. Differential dynamics of histone H3 methylation at positions K4 and K9 in the mouse zygote. *BMC Dev. Biol.* **4:** 12.

Loh Y.H., Wu Q., Chew J.L., Vega V.B., Zhang W., Chen X., Bourque G., George J., Leong B., Liu J., et al. 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38:** 431–440.

Mak W., Nesterova T.B., de Napoles M., Appanah R., Yamanaka S., Otte A.P., and Brockdorff N. 2004. Reactivation of the paternal X chromosome in early mouse embryos. *Science* **303:** 666–669.

Martin G.R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci.* **78:** 7634–7638.

Martinho R.G., Kunwar P.S., Casanova J., and Lehmann R. 2004. A noncoding RNA is required for the repression of RNApolII-dependent transcription in primordial germ cells. *Curr. Biol.* **14:** 159–165.

Matsui Y., Zsebo K., and Hogan B.L. 1992. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **70,** 841–847.

Mayer W., Niveleau A., Walter J., Fundele R., and Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature* **403:** 501–502.

McLaren A. 2003. Primordial germ cells in the mouse. *Dev. Biol.* **262:** 1–15.

McLaren A. and Lawson K.A. 2005. How is the mouse germ-cell lineage established? *Differentiation* **73:** 435–437.

McLay D.W. and Clarke H.J. 2003. Remodelling the paternal chromatin at fertilization in mammals. *Reproduction* **125:** 625–633.

Meshorer E., Yellajoshula D., George E., Scambler P.J., Brown D.T., and Misteli T. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell.* **10:** 105–116.

Mitsui K., Tokuzawa Y., Itoh H., Segawa K., Murakami M., Takahashi K., Maruyama M., Maeda M., and Yamanaka S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113:** 631–642.

Monk M., Boubelik M., and Lehnert S. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99:** 371–382.

Morgan H.D., Dean W., Coker H.A., Reik W., and Petersen-Mahrt S.K. 2004. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: Implications for epigenetic reprogramming. *J. Biol. Chem.* **279:** 52353–52360.

Morgan H.D., Santos F., Green K., Dean W., and Reik W. 2005. Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* **14:** R47–58.

Nichols J., Zevnik B., Anastassiadis K., Niwa H., Klewe-Nebenius D., Chambers I., Scholer H., and Smith A. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95:** 379–391.

Niwa H., Toyooka Y., Shimosato D., Strumpf D., Takahashi K., Yagi R., and Rossant J. 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123:** 917–929.

O'Carroll D., Erhardt S., Pagani M., Barton S.C., Surani M.A., and Jenuwein T. 2001. The polycomb-group gene Ezh2 is required for early mouse development. *Mol. Cell. Biol.* **21:** 4330–4336.

Ohinata Y., Payer B., O'Carroll D., Ancelin K., Ono Y., Sano M., Barton S.C., Obukhanych T., Nussenzweig M., Tarakhovsky A., et al. 2005. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436:** 207–213.

Okamoto I., Otte A.P., Allis C.D., Reinberg D., and Heard E. 2004. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* **303:** 644–649.

Okamoto I., Arnaud D., Le Baccon P., Otte A.P., Disteche C.M., Avner P., and Heard E. 2005. Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. *Nature* **438:** 369–373.

Olek A. and Walter J. 1997. The pre-implantation ontogeny of the H19 methylation imprint. *Nat. Genet.* **17:** 275–276.

Oswald J., Engemann S., Lane N., Mayer W., Olek A., Fundele R., Dean W., Reik W., and Walter J. 2000. Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10:** 475–478.

Peters A.H., O'Carroll D., Scherthan H., Mechtler K., Sauer S., Schofer C., Weipoltshammer K., Pagani M., Lachner M., Kohlmaier A., et al. 2001. Loss of the *Suv39h* histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107:** 323–337.

Piotrowska-Nitsche K., Perea-Gomez A., Haraguchi S., and Zernicka-Goetz M. 2005. Four-cell stage mouse blastomeres have different developmental properties. *Development* **132:** 479–490.

Reik W., Dean W., and Walter J. 2001. Epigenetic reprogramming in mammalian development. *Science* **293:** 1089–1093.

Ren B., Chee K.J., Kim T.H., and Maniatis T. 1999. PRDI-BF1/Blimp-1 repression is mediated by corepressors of the Groucho family of proteins. *Genes Dev.* **13:** 125–137.

Resnick J.L., Bixler L.S., Cheng L., and Donovan P.J. 1992. Long-term proliferation of mouse primordial germ cells in culture. *Nature* **359:** 550–551.

Rossant J. 2001. Stem cells from the mammalian blastocyst. *Stem Cells* **19:** 477–482.

Rougier N., Bourc'his D., Gomes D.M., Niveleau A., Plachot M., Paldi A., and Viegas-Pequignot E. 1998. Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12:** 2108–2113.

Roy S. and Ng T. 2004. Blimp-1 specifies neural crest and sensory neuron progenitors in the zebrafish embryo. *Curr. Biol.* **14:** 1772–1777.

Saitou M., Barton S.C., and Surani M.A. 2002. A molecular program for the specification of germ cell fate in mice. *Nature* **418:** 293–300.

Saitou M., Payer B., Lange U.C., Erhardt S., Barton S.C., and Surani M.A. 2003. Specification of germ cell fate in mice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358:** 1363–1370.

Santos F., Hendrich B., Reik W., and Dean W. 2002. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev. Biol.* **241:** 172–182.

Santos F., Peters A.H., Otte A.P., Reik W., and Dean W. 2005. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev. Biol.* **280:** 225–236.

Sarmento O.F., Digilio L.C., Wang Y., Perlin J., Herr J.C., Allis C.D., and Coonrod S.A. 2004. Dynamic alterations of specific histone modifications during early murine development. *J. Cell Sci.* **117:** 4449–4459.

Schaner C.E., Deshpande G., Schedl P.D., and Kelly W.G. 2003. A conserved chromatin architecture marks and maintains the restricted germ cell lineage in worms and flies. *Dev. Cell* **5:** 747–757.

Sciammas R. and Davis M.M. 2004. Modular nature of Blimp-1 in the regulation of gene expression during B cell maturation. *J. Immunol.* **172:** 5427–5440.

Seki Y., Hayashi K., Itoh K., Mizugaki M., Saitou M., and Matsui Y. 2005. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev. Biol.* **278:** 440–458.

Seydoux G. and Dunn M.A. 1997. Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of *Caenorhabditis elegans* and *Drosophila melanogaster*. *Development* **124:** 2191–2201.

Seydoux G. and Strome S. 1999. Launching the germline in *Caenorhabditis elegans*: Regulation of gene expression in early germ cells. *Development* **126:** 3275–3283.

Shaffer A.L., Lin K.I., Kuo T.C., Yu X., Hurt E.M., Rosenwald A., Giltnane J.M., Yang L., Zhao H., Calame K., and Staudt L.M. 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity* **17:** 51–62.

Shapiro-Shelef M., Lin K.I., McHeyzer-Williams L.J., Liao J., McHeyzer-Williams M.G., and Calame K. 2003. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity* **19:** 607–620.

Short R.V. 2000. Where do babies come from? *Nature* **403:** 705.

Smith A.G. 2001. Embryo-derived stem cells: Of mice and men. *Annu. Rev. Cell Dev. Biol.* **17:** 435–462.

Surani M.A. 2001. Reprogramming of genome function through epigenetic inheritance. *Nature* **414:** 122–128.

———. 2005. Nuclear reprogramming by human embryonic stem cells. *Cell* **122:** 653–654.

Surani M.A., Ancelin K., Hajkova P., Lange U.C., Payer B., Western P., and Saitou M. 2004. Mechanism of mouse germ cell specification: A genetic program regulating epigenetic reprogramming. *Cold Spring Harbor Symp. Quant. Biol.* **69:** 1–9.

Tada M., Tada T., Lefebvre L., Barton S.C., and Surani M.A. 1997. Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J.* **16:** 6510–6520.

Tada M., Takahama Y., Abe K., Nakatsuji N., and Tada T. 2001. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Curr. Biol.* **11:** 1553–1558.

Tada T., Tada M., Hilton K., Barton S.C., Sado T., Takagi N., and Surani M.A. 1998. Epigenotype switching of imprintable loci in embryonic germ cells. *Dev. Genes Evol.* **207:** 551–561.

Toyooka Y., Tsunekawa N., Akasu R., and Noce T. 2003. Embryonic stem cells can form germ cells in vitro. *Proc. Natl. Acad. Sci.* **100:** 11457–11462.

Turner C.A., Jr., Mack D.H., and Davis M.M. 1994. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell* **77:** 297–306.

van der Heijden G.W., Dieker J.W., Derijck A.A., Muller S., Berden J.H., Braat D.D., van der Vlag J., and de Boer P. 2005. Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122:** 1008–1022.

Van Doren M., Williamson A.L., and Lehmann, R. 1998. Regulation of zygotic gene expression in *Drosophila* primordial germ cells. *Curr. Biol.* **8** 243–246.

Vincent S.D., Dunn N.R., Sciammas R., Shapiro-Shalef M., Davis M.M., Calame K., Bikoff E.K., and Robertson E.J. 2005. The zinc finger transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse. *Development* **132:** 1315–1325.

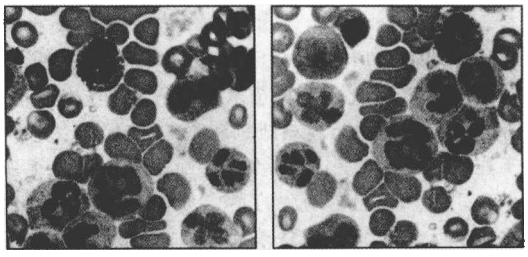
Waddington C. 1956. *Principles of embryology*. Allen & Unwin, London, United Kingdom.

Weber R.J., Pedersen R.A., Wianny F., Evans M.J., and Zernicka-Goetz M. 1999. Polarity of the mouse embryo is anticipated before implantation. *Development* **126:** 5591–5598.

Yamaguchi S., Kimura H., Tada M., Nakatsuji N., and Tada T. 2005. Nanog expression in mouse germ cell development. *Gene Expr. Patterns* **5:** 639–646.

Zhang F., Barboric M., Blackwell T.K., and Peterlin B.M. 2003. A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb. *Genes Dev.* **17:** 748–758.

# 第 21 章



# 淋巴细胞生成的表观遗传学调控

Meinrad Busslinger[1] and Alexander Tarakhovsky[2]

[1]*Research Institute of Molecular Pathology, Vienna Biocenter, A-1030 Vienna, Austria*

[2]*Laboratory of Lymphocyte Signaling, The Rockefeller University, New York, New York 10021*

## 本章目录

1. 概述
2. 淋巴细胞生成早期的世系化决定
   2.1 细胞外信号
   2.2 转录因子
   2.3 基因表达的表观遗传学调控
3. 抗原受体多样性的表观遗传学调控
   3.1 抗原受体基因重排的发育调控
   3.2 免疫球蛋白基因的亚核重定位
   3.3 免疫球蛋白基因的位点收缩
   3.4 *IgH* 和 *Igκ* 等位基因排斥的控制
4. 成熟 B 细胞的终末分化
   4.1 浆细胞分化
   4.2 成熟 B 细胞的发育可塑性
5. 结语

参考文献

## 概要

每一位免疫学家都相信一个简单的道理——淋巴细胞是研究令人费解的生命过程最好的细胞材料。细胞调控的表观遗传学机制也遵循这一规律。在发育的早期阶段，共同的淋巴源祖细胞需要决定是发育成“平凡的（boring）”B细胞还是“了不起的（terrific）”T细胞。对有兴趣研究表观遗传生理学的人来说，发育中的淋巴细胞为精确地剖析若干不连续的发育步骤提供了绝佳的机会。高效率的细胞分选手段甚至能够将发育中和成熟的淋巴细胞中各低丰度亚群大量分离获得，从而可以满足最苛求完美的技术狂。因此，淋巴细胞为鉴定那些在干细胞转化为成熟细胞的重编程过程中发挥作用的基因和调控因子提供了良机。转录因子在这个过程中起了至关重要的作用，而对“组蛋白修饰密码”的“编写者”和“阅读者”所起作用的认识也正迅速赶上。

在B细胞或T细胞的发育过程中，每一个发育中的淋巴细胞都要决定它的“社会”特征性，也就是它的抗原特异性。发育中B细胞或T细胞的唯一抗原受体的表达对于下一步筛除有害的自身免疫淋巴细胞是必需的，并促使它们发育成为能识别有限几种环境抗原分子的受体细胞。为实现这一惊人的专一性，每个B细胞或T细胞都会经历免疫球蛋白或T细胞受体基因的重排。为了实现这一过程，被染色质所保护的免疫球蛋白或T细胞受体基因位点，变得能被基因重排机器接近。被多达3Mb的DNA隔开的基因片段相互接近，然后相连接产生一个唯一性的抗原受体基因。免疫球蛋白基因重排过程包含了各种表观遗传学者熟知的风味。免疫球蛋白位点必须先转录才能发生重排的现象明显类似于*Xist*的转录，后者诱导了表观遗传学变化并导致X染色体失活。参与X染色体失活的组蛋白甲基化酶Ezh2证实了上述观点，它在免疫球蛋白基因重排的过程中也起了很重要的作用。对于那些对小RNA感兴趣的人来说，发育中的B细胞免疫球蛋白位点有双向转录现象，因此有可能产生双链小RNA，从而标记染色质使之成为重排对象。

B细胞和T细胞一旦发育成熟，就必须存活足够久以实现Warhol所谓的15分钟荣耀*，对免疫细胞来说，指的是能遭遇到抗原并对其作出响应。为了实现这一崇高的使命，B细胞和T细胞用尽从转录因子到组蛋白修饰酶等所有可能的工具。在被抗原激活后，B细胞和T细胞进入完全不同的通路。活化的T细胞分化成能产生不同细胞因子的效应细胞，这些细胞因子能辅助其他免疫细胞发挥它们的功能或是直接杀死其靶标。T细胞一直保持着其世系特征性直至死亡，与之不同的是，B细胞最终的目标就是失去其特征性，成为用于中和病原体的高度特异性的可溶性抗体的来源。为了达到这个目标，活化的B细胞会进行抗体的优化，通过突变它们的免疫球蛋白基因来提高其抗原受体对抗原的亲和力。这个过程是通过激活诱导的脱氨酶催化的，但将这个强大的“突变子”从2米长的基因组DNA中精确定位到一个微小区域的机制尚不清楚，很可能是被表观遗传学机制控制的。具高亲和力的抗原受体表达能力的B细胞再进入新一轮激活诱导的脱氨酶介导的DNA突变，使免疫球蛋白发生类别转换从而生成不同类型的抗体，从而使这些抗体具有保护生物体的多重潜力。好奇的表观遗传学家们肯定能发现这个所谓的免疫球蛋白类别转换重排中富集了许多与其他表观遗传学过程（如四膜虫DNA剔除，酵母端粒沉默和哺乳动物X染色体失活）相仿的现象。那些为优化抗体耗尽全力的B细胞迟早都会遇见抗原，这将导致重大基因组重编程，使B细胞失去其特征性，进而形成能产生抗体的浆细胞。虽然这个重要过程的转录调控机制已经相当明确，B细胞的大规模去分化机制仍对那些寻求能按意愿改造细胞特征性的人们有着巨大的吸引力。

研究淋巴细胞的分化和功能的表观遗传学机制有着决不被高估的现实意义。多种组蛋白甲基转移酶在许多淋巴癌中高表达，用专一性抑制剂将它们针对性地抑制，将有可能有助于阻止肿瘤细胞的生长和扩散。由于人们对于那些能够控制干细胞分化成淋巴细胞的因子已经有了非常全面的了解，淋巴细胞可能成为第一批能被干细胞重编程而来并实用于治疗的细胞。概括来说，对科学的热忱有助于提高免疫力，这一重要的信条使表观遗传学成为最好的免疫刺激。

---

* Andy Warhol曾预言每一个艺术家都会有15分钟的荣耀，意思相当于天生我材必有用——译者注

## 1. 概　述

组织特异性干细胞负责生命周期中整体器官，如皮肤、肠和血液系统的发育和再生。因此，干细胞具有两种独特的属性。①它们有无限的自我更新潜能，因此能在非分化状态增殖。②它们具有全能性，因此能分化成发育潜能逐步受限的各种多能性源祖细胞，产生一种器官中所有类型的细胞（图 21-1）。这些源祖细胞会继续选择某种世系，并沿着选定的通路分化成器官中具有特定功能的细胞（图 21-1）。

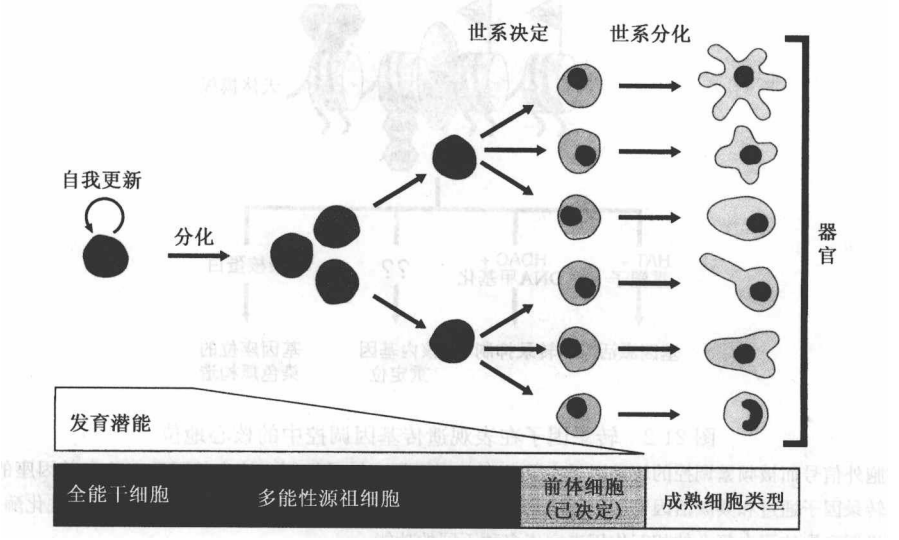



图 21-1　由单个干细胞发育成器官的示意图

器官的发育是由组织特异性的干细胞起始的，这些干细胞具有广泛的自我更新潜能和多能性，因此能产生一种器官中所有类型的细胞。干细胞首先分化成具有逐步受限潜能的多能性源祖细胞（又称转型级联放大细胞）。这些源祖细胞接下来会经历某种特定的世系化决定，然后沿着选定的路线分化成器官中具有特定功能的细胞。

多能性源祖细胞世系决定的分子水平控制，是一个重要的发育生物学问题。由于转录因子能够重编大量基因群的表达，它们在这个过程中起了十分重要的作用。为此，它们针对细胞外信号，运用多种机制来激活或抑制基因的表达（图 21-2）（Fisher 2002）。它们能通过蛋白-蛋白相互作用来拮抗其他的转录因子，从而间接影响基因的表达程序。更多情况下，转录因子通过募集具有组蛋白修饰或染色质重塑活性的共激活因子或共抑制因子来调控 DNA 元件以实现对基因转录的直接控制（图 21-2）。基因表达不仅仅由转录因子组合的存在决定，还取决于其所处的染色质环境，后者是细胞个体发育历史的反映。例如，基因活性受到 DNA 甲基化模式，染色质组蛋白修饰状态，基因相对于抑制性异染色质的核定位状态以及基因座结构的影响。

造血系统，特别是淋巴细胞的发育非常适合用于研究世系化决定和分化的表观遗传学机制，因为造血系统的世系图已非常明晰，从造血干细胞到各种成熟的血细胞之间的各个发育阶段均已被深入了解（图 21-3）（Akashi et al. 2000；Busslinger 2004）。因此，

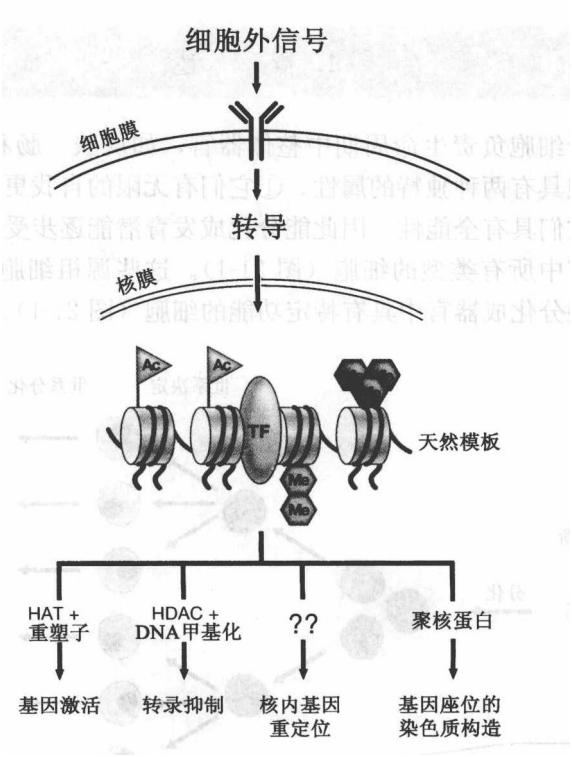



图 21-2 转录因子在表观遗传基因调控中的核心地位

根据细胞外信号而被频繁调控的转录因子主要负责基因激活、转录抑制、核内基因重定位和基因座的染色质构造。转录因子通过和共激活因子（包括组蛋白乙酰转移酶）、共阻遏因子（包括组蛋白去乙酰化酶）、染色质重塑机器和聚梳蛋白复合体相互作用来完成多种不同的功能。

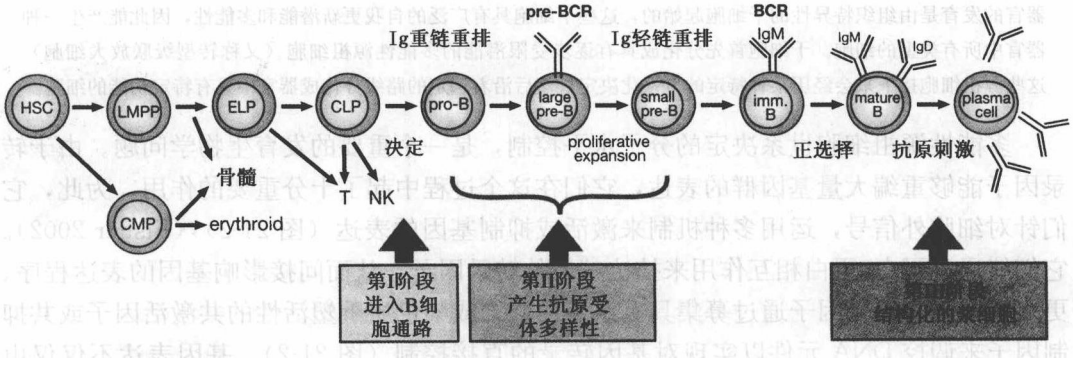



图 21-3 B细胞发育示意图

造血干细胞（HSC）通过图示的发育阶段分化成分泌免疫球蛋白的浆细胞。LMPP，预淋巴化的多能性源祖细胞；CMP，髓系共同源祖细胞；ELP，早期淋巴源祖细胞；CLP，淋巴系共同源祖细胞。I～III部分表示淋巴细胞生成的各发育阶段，这些阶段的表观遗传学机制已在本章中进行了讨论。

不同发育阶段的细胞可以通过 FACS 从造血器官中分选出来，进行体外培养，分子水平分析，经病毒表达系统改造，然后重新注入受体小鼠中，在体内研究遗传操作带来的

功能改变。虽然造血发育的表观遗传学机制只初步得到阐述，但大部分已知的表观遗传学控制的原理都适用于 B 淋巴细胞生成。

B 细胞的发育可以粗略地分为三个阶段，分别为从源祖细胞到 B 细胞的世系化决定，通过 V（D）J 重排产生抗原受体多样性以及从成熟 B 细胞最终分化成分泌免疫球蛋白的浆细胞（图 21-3）。在本章中，我们讨论控制这三个发育阶段的表观遗传学机制，这些机制同时也揭示了适用于其他发育系统的基本原理。

## 2. 淋巴细胞生成早期的世系化决定

### 2.1　细胞外信号

造血细胞的发育起始于成鼠骨髓中造血干细胞分化成髓系共同源祖细胞（common myeloid progenitor，CMP）（Akashi et al. 2000）或预淋巴化的多能性源祖细胞（lymphoid-primed multipotent progenitor，LMPP）（Adolfsson et al. 2005）。红细胞只能由 CMP 发育而来，而骨髓细胞能从 CMP 或 LMPP 发育而来。LMPP 细胞中淋巴基因 *RAG1* 和 *RAG2* 的激活标志了早期淋巴源祖细胞（earliest lymphocyte progenitors，ELP）的出现，这些淋巴源祖细胞能分化成具有 B、T 和 NK 细胞发育潜能的淋巴共同源祖细胞（common lymphoid progenitors，CLP）（图 21-3）（Busslinger 2004）

在骨髓中，细胞外信号和转录因子对于 HSC 向 CLP 和早 B 细胞（pro-B cell）源祖细胞的分化都是必须的。淋巴源祖细胞和 B 淋巴细胞的产生取决于 *c-Kit*、*Flt3* 和 *IL-7* 受体介导的信号通路（Waskow et al. 2002；Vosshenrich et al. 2003）。特别是 *Flt3* 和 *IL-7Ra* 双缺陷的小鼠完全不能产生早期 B 细胞（Vosshenrich et al. 2003）。*Flt3* 和 *IL-7R* 信号激活了与之紧密相关的转录因子 Stat5a 和 Stat5b，这两个转录因子的激活有利于淋巴源祖细胞分化成早期 B 细胞（图 21-4）（Zhang et al. 2000；Kovanen and Leonard 2004）。另外，B 细胞的发育取决于锌指转录因子 Ikaros 和 Ets 结构域蛋白 PU. 1，因为在 *Ikaros* 和 *PU. 1* 基因突变的小鼠中，不存在具有 IL-7R 标志性表达的 CLP（图 21-4）（Allman et al. 2003；Dakic et al. 2005）。

### 2.2　转录因子

淋巴源祖细胞进入 B 细胞通路取决于螺旋-环-螺旋转录因子 E2A（Tcfe2a）、早期 B 细胞因子 EBF1 和成对结构域蛋白 Pax5（BSAP）（图 21-4）。E2A 和 EBF1 共同激活 B 细胞特异性基因的表达，这为早 B 细胞的生成所必须（Busslinger 2004）。而 Pax5 缺失时，B 细胞特异性转录程序的激活不足以决定早期源祖细胞向 B 淋巴细胞发育。*Pax5* 纯合突变体中，骨髓中 B 细胞发育停滞于早期 B 细胞阶段（Nutt et al. 1997）。*Pax5*$^{-/-}$突变体中，早 B 细胞保留着类似于未决定的淋巴源祖细胞的特征性，即大量自我更新和多向发育潜能（Nutt et al. 1999）。有趣的是，在 *Pax5*$^{-/-}$源祖细胞中，髓系转录因子在淋巴组织处的异位表达能有效地诱导由淋巴向骨髓细胞的命运转化，这说明 *Pax5*$^{-/-}$突变体中早 B 细胞是一种体外可克隆的，具有骨髓分化潜能的淋巴源

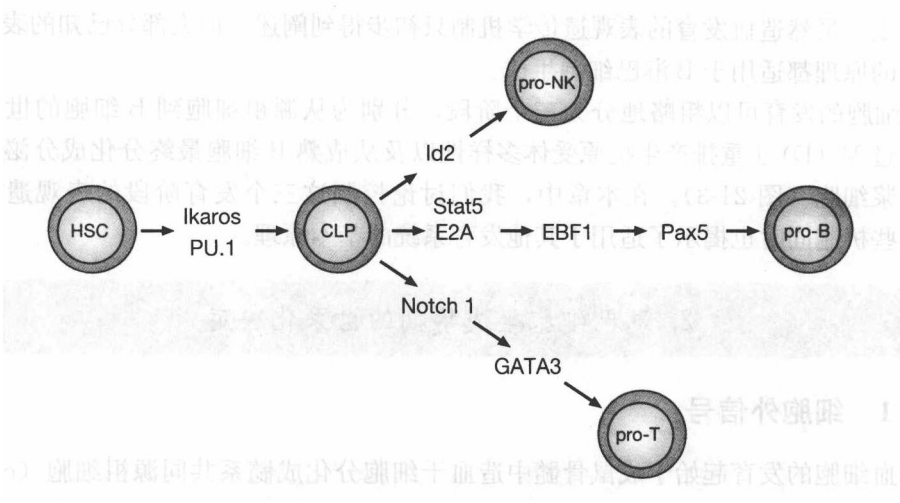



图 21-4 早期淋巴细胞形成的转录调控

在图示转录因子的调控下，造血干细胞（HSC）发育成自然杀伤细胞（NK）、B 细胞和 T 细胞世系的源祖细胞。详见 Busslinger 2004。

祖细胞（Busslinger 2004）。更重要的是，恢复 *Pax5* 基因的表达能够抑制 *Pax5*$^{-/-}$早 B 细胞的多向发育潜能，而促进其向成熟 B 细胞发育（Nutt et al. 1999）。这些实验证明，Pax5 是一个关键的 B 细胞命运决定因子，限定了淋巴源祖细胞向 B 细胞通路的发育选择。

未决定的造血源祖细胞通过“预世系化”进程，实现数种不同世系程序特异性基因的混杂表达（Miyamoto et al. 2002）。与此一致的是，*Pax5*$^{-/-}$早 B 细胞不仅表达早 B 细胞特异基因，而且还表达其他世系特异性基因（Nutt et al. 1999）。在世系决定期，Pax5 完成激活 B 细胞特异性基因表达，同时抑制世系不匹配基因表达的双重任务（图 21-5）（Nutt et al. 1999）。Pax5 激活一些编码（前）B 细胞受体（BCR）信号通路重要组分的靶基因，例如信号传递链 Igα（mb-1，Cd79a）、刺激共受体 CD19 和 CD21 以及中央接头蛋白 BLNK（图 21-5）（Busslinger 2004）。因此，Pax5 的反式激活功能促进由前 BCR 向 BCR 的信号转导，从而构成了 B 细胞发育的重要关卡。另一方面，Pax5 抑制基因则编码大量的分泌蛋白、细胞黏着分子、信号转导子和核蛋白，这些基因在红细胞、骨髓细胞和/或 T 淋巴细胞中表达（图 21-5）（Delogu et al. 2006）。其中包括 *Csf1r*（*M-CSFR*）和 *Notch1* 基因，其 Pax5 依赖性的低表达很好地阐明了决定后的 B 细胞为何不再应答骨髓细胞因子 M-CSF（Nutt et al. 1999）或诱导 T 细胞产生的 Notch 配体（Busslinger 2004）。因此，Pax5 的抑制功能通过关闭不当信号系统参与 B 细胞决定。通过与 Groucho 蛋白家族辅抑制因子的相互作用，Pax5 能够募集组蛋白去乙酰化酶（histone deacetylase，HDAC）到调控元件上，并导致基因抑制（Linderson et al. 2004）。

E2A、EBF1 以及它们的靶基因通常都在 *Pax5*$^{-/-}$早 B 细胞中表达，这表明在 B 细胞发育过程中，Pax5 作用于 E2A 和 EBF1 的下游（Nutt et al. 1997；Busslinger 2004）。与此一致的是，在 *E2A*$^{-/-}$或 *EBF1*$^{-/-}$源祖细胞中，并不存在 *Pax5* 转录物（Ikawa et al. 2004；Medina et al. 2004）。在 *E2A*$^{-/-}$源祖细胞中，*EBF1* 的表达水平很低（Ikawa et al. 2004），而在 *EBF1*$^{-/-}$源祖细胞中，*E2A* 被正常转录（Medina et al. 2004）。此

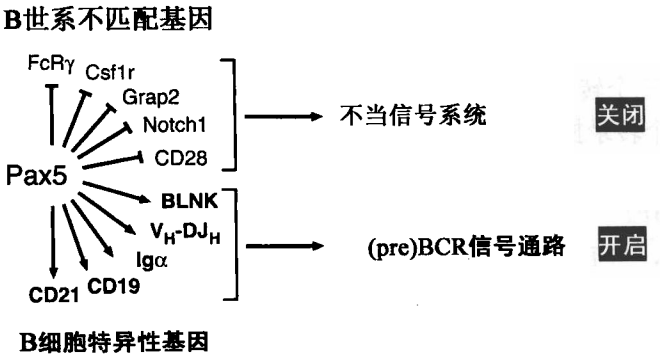



图 21-5　Pax5 在 B 细胞发育中的双重作用

Pax5 激活（pre）BCR 信号通路相关的 B 细胞特异性基因表达，抑制骨髓（*FcRγ*、*Csf1r*）或 T 淋巴细胞（*Notch1*、*CD28*、*Grap2*）中表达的不当信号转导基因。

外，通过逆转录病毒恢复 $E2A^{-/-}$ 源祖细胞中 *EBF1* 的表达能够激活 *Pax5* 转录，从而启动早 B 细胞分化（Seet et al. 2004）。因此，这三个转录因子通过 E2A-EBF1-Pax5 的遗传级联促进早期 B 细胞发育。而在 B 淋巴细胞世系中，*EBF1* 的表达则通过 Pax5 的反馈调控得以维持（图 21-4）（Horcher et al. 2001；Fuxa et al. 2004）。

## 2.3　基因表达的表观遗传学调控

在多细胞真核生物的任何一个细胞中，大部分基因处于转录抑制状态，且只在特定的发育过程中被选择性地激活。基因抑制被多重表观遗传学机制所介导，包括 DNA 甲基化、组蛋白修饰、核小体定位和染色质高级结构（Smale 2003）。对 B 细胞特异基因

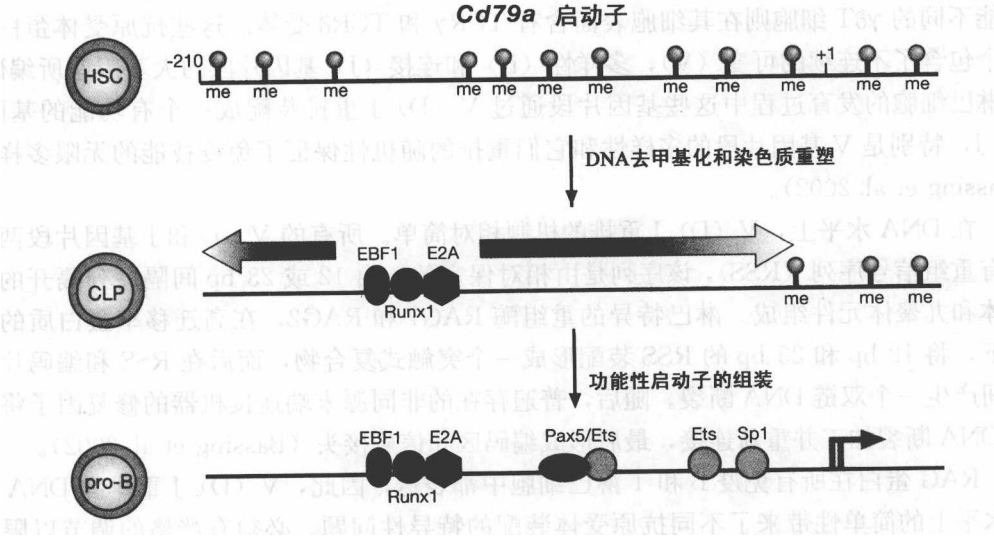



图 21-6　早期淋巴细胞生成过程中 *Cd79a* 基因的表观遗传激活

*Cd79a*（*mb-1*，*Iga*）启动子示意图，同时显示了 CpG 甲基化（me）模式和从 HSC 向早 B 细胞分化过程中不同转录因子的相继结合。

*Cd79a* 的研究为 B 细胞发育开始时基因抑制是如何被解除的问题提供了第一手资料（mb-1，Iga；Maier et al. 2004）。*Cd79a* 基因启动子除了包括 Ets 和 Runx 的结合位点之外，还包括了三个转录因子 E2A、EBF1 和 Pax5 的功能识别序列（图 21-6）（Maier et al. 2004）。这个转录控制区域的 CpG 二核甘酸在 HSC 中是被甲基化的，在 CLP 中部分去甲基化，在决定了的早 B 细胞中完全去甲基化（Maier et al. 2004）。*Cd79a* 启动子在 *E2A*$^{-/-}$ 或 *EBF1* $^{-/-}$源祖细胞中也是完全甲基化的，但能在 EBF1 和 Runx1 的协同作用下去甲基化，与 Pax5 一起作用，导致 *mb-1* 基因在决定了的早 B 细胞中被转录激活（图 21-6）（Maier et al. 2004）。因此，EBF1，可能还有 E2A，作为 B 细胞早期发育的“先锋”因子，通过启动基因激活所需的表观遗传变化，促进早期源祖细胞分化成决定的早 B 细胞。虽然组蛋白修饰和染色质结构的变化尚未得到研究，但 E2A 和 Pax5 很可能通过与 SAGA 和 p300/CBP 组蛋白乙酰转移酶（HAT）复合物的相互作用参与了开放的乙酰化染色质的形成（Massari et al. 1999；Barlev et al. 2003）。

## 3. 抗原受体多样性的表观遗传学调控

### 3.1 抗原受体基因重排的发育调控

获得性免疫系统的指导原则是每一个新产生的淋巴细胞只识别一种特异性的抗原，这样淋巴细胞的整体多样性就足以对付各种可能的抗原。为了达到这个目的，B 和 T 细胞表达世系特异性的抗原受体，以介导抗体依赖性的体液免疫或细胞免疫。BCR 包括免疫球蛋白重链（IgH）和一个 Igκ 轻链或 Igλ 轻链（IgL）。T 细胞中的 αβ 世系，组成了小鼠和人的大部分 T 淋巴细胞，它们表达 T 细胞受体（TCR）β 多肽和 TCRα，而功能不同的 γδT 细胞则在其细胞表面含有 TCRγ 和 TCRδ 受体。这些抗原受体蛋白由一个包含了不连续的可变（V），多样性（D）和连接（J）基因片段的大基因座所编码，在淋巴细胞的发育过程中这些基因片段通过 V（D）J 重排装配成一个有功能的基因。D、J，特别是 V 基因片段的多样性和它们重排的随机性保证了免疫技能的无限多样性（Bassing et al. 2002）。

在 DNA 水平上，V（D）J 重排的机制相对简单。所有的 V、D 和 J 基因片段两端都有重组信号序列（RSS），该序列是由相对保守的，由 12 或 23 bp 间隔序列隔开的七聚体和九聚体元件组成。淋巴特异的重组酶 RAG1 和 RAG2，在高迁移率蛋白质的辅助下，将 12 bp 和 23 bp 的 RSS 装配形成一个突触式复合物，而后在 RSS 和编码片段之间产生一个双链 DNA 断裂。随后，普遍存在的非同源末端连接机器的修复因子将这些 DNA 断裂加工并重新连接，最后形成编码区和信号接头（Bassing et al. 2002）。

RAG 蛋白在所有免疫 B 和 T 淋巴细胞中都表达，因此，V（D）J 重排在 DNA 模板水平上的简单性带来了不同抗原受体装配的特异性问题。必须有严格的调节以限制 RAG 蛋白只能特异性地作用于所有重排底物中的一部分（Yancopoulos and Alt 1985；Stanhope-Baker et al. 1996）。V（D）J 重排以世系和阶段特异的方式被严格调控。在 B 淋巴世系中，早 B 细胞中 *IgH* 基因的重排先于前 B 细胞中 Igκ 和 Igλ 的重排，而 *TCR*β

和 *TCRα* 基因的重排则分别发生在早 T 细胞和前 T 细胞中。此外，*IgH* 基因 V（D）J 重排具有一个明确的时间顺序，即 $D_H$-$J_H$重排优先于 $V_H$-$DJ_H$重排。在早 T 细胞发育过程中，TCRβ 基因座的重排以同样的顺序（$D_\beta$-$J_\beta$在 $V_\beta$-$DJ_\beta$之前）进行。因此，必须有控制机制的存在以保护所有的 V 基因在 D-J 重排时不被 RAG 剪切，并在 V-DJ 重排时保证在上百个 V 基因中只有一个发生重排。因此，抗原受体产生的过程完全依赖于对 RAG1/2 重组酶与 RSS 作用的精确调控。

成功的 *IgH* 或 *TCRβ* 基因的 V-DJ 重排导致 IgH 或 TCRβ 蛋白作为前 BCR 或前 TCR 复合物组分的表达，这作为一个重要的检验点，可以使第二个已发生 DJ 重排的等位基因不再进行 V-DJ 重排并促进前 B 和前 T 细胞的发育，引发 *IgL* 或 *TCRα* 基因的重排。最终，具有信号功能的 BCR 或 TCR 的表达，在未成熟 B 或 T 细胞中，通过抑制 RAG1/2 基因的转录，完全阻断 V（D）J 重排（Jankovic et al. 2004）。然而，自身免疫 BCR 产生的信号能重新启动免疫球蛋白轻链基因的重排，以产生一个具有新抗原特异性的 BCR（受体编辑；Jankovic et al. 2004）。另外，细胞因子 IL-7 产生的信号为促进早 T 细胞中 TCRγ 基因的重排所必需（Schlissel et al. 2000）。因此，V（D）J 重排不仅受其内在的发育和世系特异性的核机制所调控，还受产生于细胞表面的外来信号的调控。

V（D）J 重排的发育和位点特异性限制主要是发生于表观遗传水平（Krangel 2003）。在非淋巴细胞中，*Ig* 和 *TCR* 基因位于难接近的染色质区域上，因为在肾细胞中外源表达的 RAG 蛋白能有效切割转染的游离重组底物，却不能作用于内源的抗原受体基因（Romanow et al. 2000）。此外，在提取的淋巴细胞核中添加重组 RAG 蛋白，只能对那些用于提核的细胞中正在活跃发生 V（D）J 重排的 *Ig* 或 *TCR* 基因进行切割（Stanhope-Baker et al. 1996）。因此，基因重排的世系特异性和时间顺序是由连续的局部染色质的开放所决定，以允许特定的 RSS 被 V（D）J 重排酶所作用。染色质既保护 RSS 又引导对它们的切割，表明了一个“染色质密码”的存在，这个“染色质密码”标记了重组位点和/或促进 RAG 介导的切割。

组蛋白赖氨酸残基的乙酰化不仅是开放染色质的典型特征，而且在决定 *Ig* 和 *TCR* 基因座的染色质可接近性方面起了重要作用，因为它能够划分可重组的基因片段区域（McMurry and Krangel 2000；Chowdhury and Sen 2001）。对组蛋白乙酰化状态的分析揭示了 *IgH* 基因座不同染色质区域逐步激活的现象（Chowdhury and Sen 2001）．一个包含 $D_H$、$J_H$和 $C_\mu$基因片段的 120Kb 的基因组区域在 V（D）J 重排之前首先发生高乙酰化。$D_H$-$J_H$重排随后诱导 $D_H$近端的 $V_H$基因发生组蛋白乙酰化和重排（Chowdhury and Sen 2001）。最后，包括大部分 $V_H$基因的远端的 2Mb 区域被 IL-7 信号激活（Chowdhury and Sen 2001）。对发育中 T 细胞的 *TCRα/δ* 基因座的详细分析也显示组蛋白 H3 超乙酰化区域和 V（D）J 重组酶可接近区域是完全重叠的（McMurry and Krangel 2000）。因此，组蛋白乙酰化似乎是控制重组起始及进程的染色质修饰模式的必需组分（Krangel 2003）。

但是，组蛋白去乙酰化酶抑制剂对体内 V（D）J 重排几乎没有影响，因此，乙酰化本身并不足以促进重组的发生（McBland and Boyes 2000）。此外，在 pro-B 细胞中可

以惊人地同时存在组蛋白 H3 高乙酰化和 V（D）J 低重排（Hesslein et al. 2003；Su et al. 2003）。在缺失能三甲基化组蛋白 H3 第 27 位赖氨酸（H3K27me3）的组蛋白赖氨酸甲基化酶（HKMT）Ezh2 的早 B 细胞中，$D_H$末端 $V_H$基因簇上正常的组蛋白乙酰化水平并不能维持末端 $V_H$-$DJ_H$的重排（Su et al. 2003）。高水平 H3K27me3 与远端而不是近端 $V_H$基因相关联的发现表明了 H3K27 甲基化在 $V_H$基因重排中的区域特异性作用（Su et al. 2003）。在野生型和 *Ezh-2* 缺陷型的早 T 细胞中，*TCRβ* 基因的重组效率相同，这一现象凸显了 Ezh2 介导的调节作用对 *IgH* 基因座的选择性（Su et al. 2005）。因此，其他染色质修饰很可能参与了早 B 细胞中近端 $V_H$基因 V（D）J 重排和早 T 细胞中 *TCR*β 重排的调控。组蛋白 H3 第 4 位赖氨酸的双甲基化（H3K4me2）是一个激活型的组蛋白标记，它也与 *IgH* 和 *TCR*β 基因片段的可接近程度呈相关性（Morshead et al. 2003）。相反，H3 第 9 位赖氨酸的双甲基化（H3K9me2）是一个抑制型的染色质标记，与 *IgH* 和 *TCR*β 基因片段的 V（D）J 重排呈负相关性（Morshead et al. 2003）。将 H3K9 HKMT G9a 锁定于 *TCR*β 微基因座的种系 PDβ1 启动子，导致其周围的染色质不可接近从而防止 V（D）J 重排，这一最新研究成果证实了 H3K9me2 对重排的抑制功能（Osipovich et al. 2004）。

有利于 V（D）J 重组酶接近的组蛋白修饰模式必然是，通过发生在抗原受体基因上的事件，在重排之前得以建立。在组蛋白修饰图可被实验绘制之前，人们就已经知道，尚未重排的基因片段中短的正义链 RNA 的转录发生于 V（D）J 重排之前（Yancopoulos and Alt 1985）。位于抗原受体基因的增强子和启动子为 V（D）J 重排所必需，这一发现进一步支持了转录在控制位点可接近性方面的潜在作用。内源增强子和启动子的缺失导致抗原受体位点 V（D）J 重排的减少或消失，而额外的世系特异性增强子的引入则会产生新的 V（D）J 重排模式（Bassing et al. 2002；Krangel 2003）。V、D、J 片段携带的大量启动子控制相对短距离的邻近序列的重排，而增强子则对 V（D）J 重排进行长距离调控（Bassing et al. 2002；Krangel 2003）。启动子上预起始复合体的组装可能引发该处核小体的解离，这可以促进重组酶的接近，即便组蛋白修饰不发生改变。但更有可能的是，启动子主动参与形成一个允许重组的染色质结构，因为延伸阶段的 RNA 聚合酶 II 复合物自己携带着组蛋白乙酰转移酶，可能帮助组蛋白乙酰化沿着转录区域蔓延（Orphanides and Reinberg 2000）。基因转录也导致复制依赖型组蛋白 H3 被变体 H3.3 置换，这也被认为有利于维持转录区域染色质的可接近状态（Chow et al. 2005；Mito et al. 2005）。

如上所述，每一个抗原受体位点都包括成百上千的 RSS，虽然在任一淋巴细胞中，只有少数几个 RSS 会在特定的发育阶段被切割。因此可以想象，每一 RSS 位点 DNA 序列的差异有可能会影响它们的切割效率。对人造 V（D）J 重排底物的分析的确说明在 RSS 七聚体和九聚体元件，以及不太保守的间隔序列和相邻编码序列中出现的天然差异能够影响重组频率，因此对抗原受体所有 V、D、J 基因片段中特定成员的使用差异起一定的决定作用（Lee et al. 2003）。在以“组蛋白密码”学说为核心的模型框架中，选择性切割的实现，应该就是将特定 RSS 或其邻近序列的特征性表述为特定的组蛋白修饰模式，从而标记出 RAG 介导的切割位点的过程。这个密码可能是在反义转录

物的帮助下建立的。确实，在早 B 细胞中，涵盖整个 $V_H$基因簇的基因间隔区反义转录发生于 *IgH* 基因座的 $V_H$-$DJ_H$重排之前（Bolland et al. 2004）。这些长的反义转录物可能介导了染色质重塑，以便在重排之前打开大的 $V_H$基因区域。或者，这些反义转录物能与短的正义种系 $V_H$转录物形成杂合双链 RNA，然后被 RNA 干扰机器加工成 siRNA，从而将 HKMT 募集到重排位点（Bolland et al. 2004；关于 RNAi 机器详见第 8 章）。作为这个假设的延伸，我们猜想特定 RSS 位点特异性的正义链种系转录物可能会产生双链 RNA，将组蛋白修饰酶定位到这一而非其他 RSS 序列上。如果这一猜想能被实验证明，这个假定的机制就能够解释 RAG 介导的特定 RSS 位点切割的精确性和选择性。有趣的是，最近有研究表明，RAG2 蛋白能够直接与组蛋白相互作用，因此，可能在解读特定 RSS 序列中组蛋白修饰模式上起重要作用（West et al. 2005）。

## 3.2　免疫球蛋白基因的亚核重定位

核外周和着丝粒周围的异染色质是细胞核中两个主要的抑制性空间，它们对于造血细胞中非活化状态基因的扩展具重要意义（Brown et al. 1997；Baxter et al. 2002）。根据基因的活跃状态，它们被重新定位于这些抑制性空间和促进基因转录的核中心区域（Brown et al. 1997；Baster et al. 2002）。有趣的是，在所有早 B 细胞，包括非决定的淋巴源祖细胞中，*IgH* 和 *Igκ* 基因座总是默认存在于核外周（Kosak et al. 2002）。*IgH* 基因座通过远端 $V_H$基因锚定于核外周，通过近端 *IgH* 区域靠向促进 $D_H$-$J_H$重排的核中心区域（Fuxa et al. 2004）。在 B 细胞发育的起始阶段，*IgH* 基因座激活的引发步骤包括 *IgH* 和 *Igκ* 基因座从核外周向核中心区域的重定位（Kosak et al. 2002）。这个亚核重定位可能有利于染色质展开和种系转录，致使近端 $V_H$-$DJ_H$重排。有微弱的证据认为，EBF1 和 Pax5 分别在 *IgH* 和 *Igκ* 基因座的中心重定位过程中起作用（Fuxa et al. 2004；Sato et al. 2004）。

虽然在早 B 细胞中，*IgH* 和 *Igκ* 基因座的等位基因都重新定位于中心核区域（Kosak et al. 2002；Fuxa et al. 2004），但是在成功进行了 V（D）J 重排的成熟 B 细胞中，这两个等位基因具有不同的表现（Skok et al. 2001）。在 B 细胞激活后，成功重排的 *Ig* 等位基因仍远离着丝粒群，增强其表达（Skok et al. 2001）。同时，无效的 *IgH* 和 *Igκ* 等位基因在 B 细胞激活后重新定位并沉默于着丝粒异染色质区。有趣的是，着丝粒区对无效 *IgH* 和 *Igκ* 基因座的募集是通过它们的远端 V 基因区域发生的，这表明相同的 DNA 序列既能参与核外周也能参与着丝粒异染色质区域对沉默 *Ig* 基因座的募集（Roldan et al. 2005）。

## 3.3　免疫球蛋白基因的位点收缩

*IgH* 基因座中约 200 个 $V_H$基因覆盖了 2.4Mb 的区域，并依其序列相似性和相对于近端 $D_H$基因片段的位置被分为 15 个远端、中心或近端 $V_H$基因家族。在非 B 淋巴细胞和淋巴源祖细胞中，两个 *IgH* 等位基因在核外周呈现一种伸展的构象（Kosak et al. 2002）。相反，在已决定的 B 细胞中，*IgH* 基因座经过长距离收缩，将远端 $V_H$基因置于已重排的近端 $DJ_H$区域附近，从而促进 $V_H$-$DJ_H$的重排（图 21-7）（Kosak et

al. 2002；Fuxa et al. 2004）。*Igκ* 基因座的约 140 个 Vκ 基因也覆盖了 3Mb 的区域，与 *IgH* 的长度相仿。与早 B 细胞中的 *IgH* 基因相似，小前 B 细胞和未成熟 B 细胞中的 *Igκ* 基因座都经历了收缩的过程，这就说明，在重排的细胞中两个 *Ig* 基因座都处于一个收缩状态（Roldan et al. 2005）。用远端、中心和近端的基因探针进行荧光原位杂交分析表明各 *Ig* 子区域的环化导致了 *IgH* 和 *Igκ* 基因座的长距离收缩（图 21-7）（Roldan et al. 2005）。

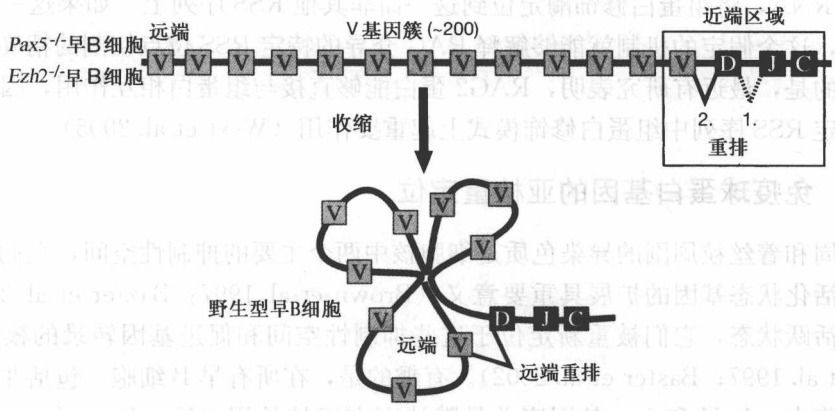



图 21-7 早 B 细胞中免疫球蛋白重链基因座的收缩

*IgH* 基因座包括一个含多样性区（D）、连接区（J）、恒定区（C）基因片段的近端区域和一个大的可变区（V）基因簇，包括约 200 个基因覆盖了约 2.4Mb 的区域。在 $Pax5^{-/-}$ 或 $Ezh2^{-/-}$ 早 B 细胞中，*IgH* 基因座以一种伸展的构象存在，使 V（D）J 重排只能发生在近端区域。在野生型早 B 细胞中，由于 *IgH* 基因座通过环化发生收缩，所有的 $V_H$基因都可以参与 $V_H$-$DJ_H$重排。

尽管 $V_H$基因沿整个 $V_H$基因簇包括最远端的 $V_H$J558 家族都以高乙酰化染色质状态存在且易被接近（Hesslein et al. 2003），但远端 $V_H$-$DJ_H$重排不发生于 $Pax5^{-/-}$ 早 B 细胞中（Nutt et al. 1997）。远端 $V_H$-$DJ_H$重排的失败与 *IgH* 基因座收缩的缺失并存（图 21-7），且能通过在 $Pax5^{-/-}$ 早 B 细胞中以逆病毒表达 Pax5 来恢复（Fuxa et al. 2004）。因此，Pax5 是早 B 细胞中 *IgH* 基因座收缩的关键调节蛋白。组蛋白甲基转移酶 Ezh2 也被认为在 *IgH* 基因座的收缩过程中起作用，因为 HSC 中 *Ezh2* 的条件性失活导致远端 $V_H$-$DJ_H$重排的减弱，尽管在 Ezh2 缺陷的早 B 细胞中远端 $V_H$基因仍保持充分的染色质可接近性（Su et al. 2003）。有趣的是，尽管 *Pax5* 和 *Ezh2* 缺陷的成熟早 B 细胞中有相似的 *IgH* 重排表型，但是它们没有遗传相关性（Fuxa et al. 2004）。因此 Pax5 有可能作为一个序列特异的靶定因子将包含 Ezh2 的 Polycomb 抑制复合物 2（PRC2）募集到 *IgH* 基因座上的选定区域。由此而产生的组蛋白 H3 第 27 位赖氨酸甲基化（H3K27me3）可能吸引 PRC1 复合物并诱导靶定区域染色质的收缩（Francis et al. 2004；在第 11 章讨论），因此产生 *IgH* 基因座的环化和收缩。或者，基因座收缩可能不取决于核中组蛋白的修饰，而是需要细胞质中含 Ezh2 的甲基转移酶复合物来催化信号蛋白的赖氨酸甲基化，这已知能通过与 GTP/GDP 置换因子 VavI 的结合来调节肌动蛋白的多聚化（Su et al. 2005）。

### 3.4　*IgH* 和 *Igκ* 等位基因排斥的控制

等位排斥保证了两个 *Ig* 等位基因中只有一个发生重排，从而导致 B 细胞只表达具有独特抗原特异性的单个抗体分子。等位基因排斥的过程被分为两个不同的阶段。在引发阶段，两 *Ig* 等位基因中的一个被选中进行不同的表观遗传学标记并首先发生重排，这排除了两个等位基因同时发生重组的可能性。随后，已重排的等位基因的表达会通过反馈抑制阻止第二个等位基因的重组，由此维持等位基因排斥。在发育早期的着床阶段，等位基因排斥已经被引发，此时，每个细胞中抗原受体基因的两个等位基因开始发生不同步复制（Mostoslavsky et al. 2001）。通过目前尚未知的表观遗传机制，被随机选中优先复制的母系或父系来源的 *Ig* 基因，几乎总是在未成熟 B 淋巴细胞中先发生重排（Mostoslavsky et al. 2001）。*IgH* 第二次发生的 $V_H$-$DJ_H$ 重排是受调控的阶段，因为在早 B 细胞发育中两个 *IgH* 等位基因都能发生 $D_H$-$J_H$ 重排（Bassing et al. 2002）。然而，等位基因特异性的表观遗传学标记（在早期胚胎中建立）是如何被翻译成两个 *IgH* 等位基因相继发生 $V_H$-$DJ_H$ 重排的激活信号呢？人们对此仍毫无所知。一个 *IgH* 等位基因的成功重排导致细胞表面前 B 细胞受体（前 BCR）Ig$\mu$ 蛋白的表达。这个受体是以下过程的关键检验点，它促进大前 B 细胞的大量增殖，随后诱导其分化为小前 B 细胞，并维持已发生 $DJ_H$ 重排的 *IgH* 基因的等位基因排斥（Kitamura and Rajewsky 1992；Bassing et al. 2002）。随着前 BCR 信号的产生，RAG 蛋白的表达迅速消失（图 21-8），这终止了所有 V（D）J 重排的继续进行，为在大前 B 细胞中建立等位基因排斥提供了条件（Grawunder et al. 1995）。前 BCR 信号还导致组蛋白去乙酰化并由此降低了小前 B 细胞中 $V_H$ 基因的可接近性，这可能是巩固等位基因排斥的反馈调节机制（Chowdhury and Sen 2003）。然而，应答于前 BCR 信号而发生的 *IgH* 收缩的迅速逆转提供了一个更为可能的机制，它将 $V_H$ 基因与近端 *IgH* 区域在空间上分隔开（图 21-8），进而阻止已发生 $DJ_H$ 重排的第二个 *IgH* 等位基因再度发生 $V_H$-$DJ_H$ 重排（Roldan et al. 2005）。此外，前 BCR 信号也能够引导非功能 *IgH* 等位基因迅速重新定位到抑制性的着丝粒区域（Roldan et al. 2005）。因此，在前 BCR 信号通路中短暂的无 RAG 时间窗口内，基因座去收缩并被招募到着丝粒区域改变了已发生 $DJ_H$ 重排的 *IgH* 等位基因，这样即便小前 B 细胞中 RAG 蛋白随后再次表达，$V_H$-$DJ_H$ 重排也不能够再次发生（图 21-8）。

通过利用甲基化敏感的限制酶检测 DNA 甲基化模式（Mostoslavsky et al. 1998；Goldmit et al. 2002；2005）并在内源 *Igκ* 的 $J_κ1$ 元件中插入了一个 *GFP* 基因的杂合子 $κ^o$-GFP 报告小鼠的分析（Liang et al. 2004），对 *Igκ* 等位基因排斥的起始进行了大量的研究。在所有的非 B 和早 B 细胞中 *Igκ* 的 CpG 二核苷酸都是高甲基化的，但在前 B 细胞中有且只有一个等位基因特异性地去甲基化（图 21-9）（Mostoslavsky et al. 1998；Liang et al. 2004）。这一单等位基因去甲基化现象发生在重排之前，并且同时依赖于基因内含子中的增强子和 3′端 κ 增强子的活性（Mostoslavsky et al. 1998）。去甲基化的等位基因处于可接近的染色质状态，因为它表现出 DNase I 敏感，组蛋白 H3 和 H4 高乙酰化，并在早 B 细胞中远离着丝粒异染色质区域（图 21-9）（Goldmit et al. 2002，2005）。因此，只有非甲基化的 *Igκ* 等位基因启动种系转录和 $V_κ$-$J_κ$ 重排（Goldmit et

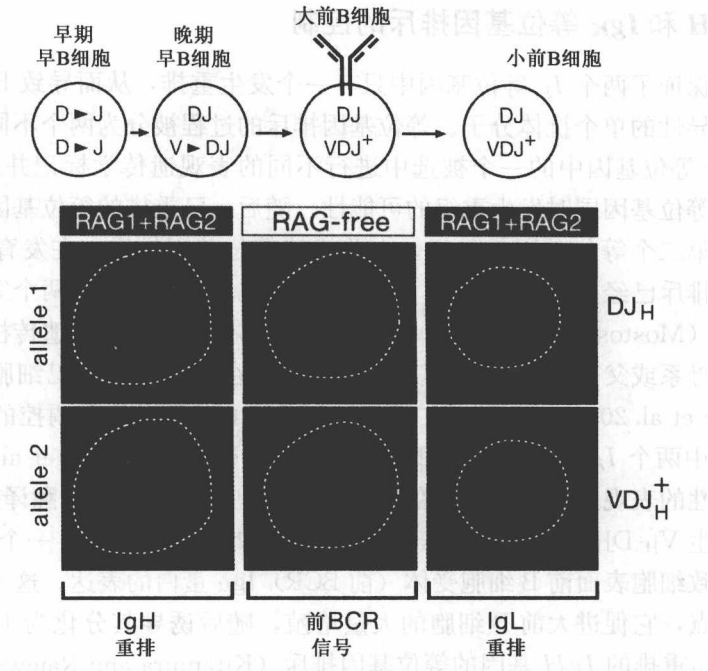



图 21-8 前 B 细胞中 *IgH* 基因座去收缩引发等位基因排斥

在早期早 B 细胞中，$D_H$-$J_H$重排同时发生在两个 *IgH* 等位基因上，而在晚期早 B 细胞中只有一个等位基因会经历 $V_H$-$DJ_H$重排。用三维 DNA-FISH 的方法，利用分别标记 *IgH* 基因座远端（红色）和近端（绿色）区域的荧光探针分析早 B 和前 B 细胞的细胞核。同一细胞中的两个 *IgH* 等位基因显示在两个代表性的共聚焦平面上。前 BCR 信号不仅导致 RAG 蛋白的迅速减少，而且导致了 *IgH* 基因座的去收缩。虽然两个等位基因都发生了去收缩，但只有未完成 DJH 重排的那个 *IgH* 等位基因才会完全伸展。具功能的重排了的等位基因（$VDJ^+$）上的两个信号由于间隔 DNA 序列的消失而隔得更近一些。FISH 数据来自 Roldan 等 2005。

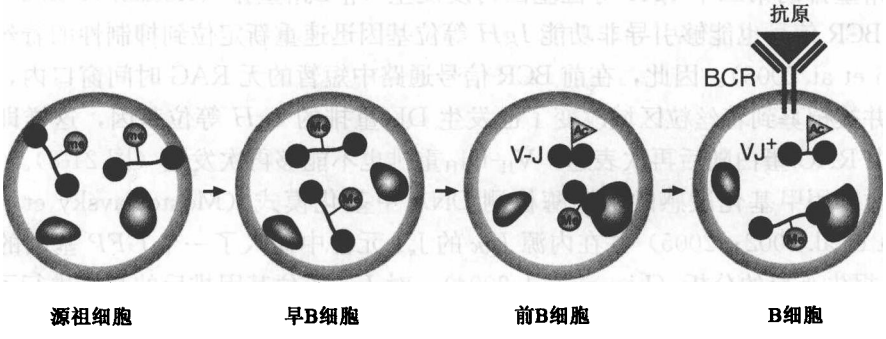



图 21-9 *Igκ* 基因座等位基因排斥的调控机制

亚核重定位、DNA 去甲基化和组蛋白乙酰化共同参与了早 B 细胞中发生 $V_κ$-$J_κ$重排的 *Igκ* 等位基因的选择。图中标记了 *Igκ* 基因座的远端 $V_κ$区域（红色）和近端 $J_κ$-$C_κ$功能域（绿色）以及相对应的核外周（灰色）和着丝粒异染色质（蓝色）等抑制性区域的定位。在包括激活的成熟 B 细胞在内的不同发育阶段中，两个 *Igκ* 等位基因的基因座收缩、DNA 甲基化（me）和组蛋白乙酰化（ac）状态。

al. 2002；Liang et al. 2004），而在小前 B 细胞中两个等位基因都发生基因座收缩（图 21-9）(Roldan et al. 2005)。令人意外的是，前 B 细胞中第二个 *Igκ* 等位基因像 *IgH* 一样（Roldan et al. 2005）被重新定位到着丝粒异染色质区域（Goldmit et al. 2005）。这种单等位基因着丝粒招募现象（图 21-9）也许解释了为什么 DNA 甲基化的等位基因缺少组蛋白乙酰化，并结合有 HP1γ 和 Ikaros 蛋白，它们与组蛋白去乙酰化酶复合物一起富集于着丝粒异染色区域（Goldmit et al. 2005）。有趣的是，被重新定位于着丝粒簇的 *Igκ* 等位基因复制较晚，这与已知的早期胚胎中前 B 细胞中 *Igκ* 单等位基因重排起始导致的非同步复制模式相一致（Mostoslavsky et al. 2001）。奇怪的是，在 $\kappa^{o}$-GFP 报告小鼠的前 B 细胞中只有很小一部分（5%）发生 *Igκ* 基因座激活（Liang et al. 2004）。基于这个结果，人们推测在前 B 细胞中与 *Igκ* 顺式调控元件相结合的某些转录因子的量非常有限，这些因子共同结合到 *Igκ* 增强子上亦属罕见，只随机的发生在其中一个等位基因中。因此，等位基因竞争有限的转录因子而产生的随机增强子激活可能参与了 *Igκ* 的等位基因排斥（Liang et al. 2004）。一个 *Igκ* 等位基因的成功重排导致了细胞表面 BCR 的表达，它随即通过抑制 RAG1/2 重组酶的表达来维持第二个 *Igκ* 等位基因的等位排斥现象（Jankovic et al. 2004）。

## 4. 成熟 B 细胞的终末分化

### 4.1　浆细胞分化

V（D）J 重排的完成和 B 细胞表面免疫球蛋白（Ig）受体的表达标志着 B 淋巴细胞中发生抗原非依赖性阶段的结束。此后，B 细胞的命运开始取决于抗原诱导的受体信号（Rajewsky 1996）。在没有抗原的时候，外周 B 细胞保持着一种静息状态，它们的生存依赖于细胞表面 BCR 的滋养信号（Kraus et al. 2004）。对静息 B 细胞和活跃 B 细胞的染色质构成进行比较的结果表明静息 B 细胞整体组蛋白甲基化水平较低（Baxter et al. 2004）。静息 B 细胞在细胞激活过程中，其相对较高的组蛋白乙酰化水平处于稳定状态（Baxter et al. 2004）。组蛋白赖氨酸甲基化水平的整体降低，包括组蛋白 H3K9 甲基化的缺失，与静息 B 细胞中其他组成性异染色质特征（如 Ikaros 和 HP1 的结合）的缺失相关（Baxter et al. 2004）。B 细胞的激活使组蛋白的甲基化水平得以恢复，进而导致 B 细胞功能所必须的基因（*Pax5*）上激活型的 H3K4me3 标记的增加，同时也导致沉默基因（*Dntt*）上抑制型的 H3K9me2 修饰的增加（Baxter et al. 2004）。

在免疫应答过程中，激活诱导的染色质重塑也许有助于保持 B 细胞的特征性。但是，B 细胞基因组必须能够应答于抗原诱导的重编程，因为激活的 B 细胞在遭遇抗原之后能够直接分化成产生抗体的浆细胞（Calame et al. 2003）。另外，抗原刺激还能够引发生发中心反应。在这个反应中，成熟的 $IgM^{low}IgD^{high}$ B 细胞发生了免疫球蛋白类型的转换，并且在活化诱导的胞嘧啶核苷脱氨酶（AID；Honjo et al. 2004）的帮助下突变它们的 *Ig* 基因。由这些过程产生的 Ig 蛋白完美地记忆 B 细胞的分化和维持或是浆细胞

的发育，后者能够产生针对一种特定抗原的高亲和力抗体（Honjo et al. 2004）。

生发中心反应时间的选择和 B 细胞向浆细胞的转变这两个过程由两个相互排斥的转录抑制蛋白——Bcl6 和 Blimp1 所调节（图 21-10）（Turner et al. 1994；Ye et al. 1997）。Bcl6 在成熟初始 B 细胞中表达水平低，但一些 B 细胞在抗原刺激之后，Bcl6 的表达水平迅速上调（Fukuda et al. 1997）。在遭遇抗原后 Bcl6 表达水平不发生上调的细胞分化成浆细胞，作为产生低亲和力抗体的最初来源（Fukuda et al. 1997）。相反，Bcl6 表达水平发生上调的 B 细胞进入生发中心反应，通过 Bcl-6 介导的控制浆细胞分化的基因的抑制来维持它们 B 细胞的特征（Shaffer et al. 2000）。Bcl6 的主要靶位点是 *Prdm1* 基因，它能够编码转录因子 Blimp1（图 21-10）。有趣的是，Pax5 似乎辅助了 Bcl6 对 *Blimp1*（*Prdm1*）基因的抑制（图 21-10）（Delogu et al. 2006）。但是，一旦 Blimp1 表达，它会抑制 B 细胞的转录程序，包括 Bcl6 和 Pax5 的表达，同时诱导浆细胞特异基因的转录（Shaffer et al. 2002；Calame et al. 2003）。

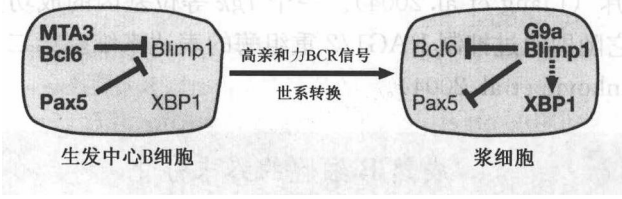



图 21-10　转录抑制决定 B 细胞和浆细胞的命运

Bcl6 和 Blimp1 通过一系列的抑制性手段抑制基因转录。Bcl6 的一个特征是利用组蛋白修饰以外的赖氨酸乙酰化来控制基因抑制（Bereshchenko et al. 2002）。Bcl6 能够与 MTA3 蛋白相互作用，后者是辅抑制蛋白复合物 Mi-2/NuRD 中的一个组分，且在生发中心高表达（Fujita et al. 2004）。RNAi 介导的 MTA3 的缺失导致 B 细胞中 Bcl-6 抑制的靶基因的重激活，因此，Bcl6 与包含 MTA3 的 Mi-2/NuRD 复合物的相互作用对于基因抑制是必须的（Fujita et al. 2004）。Bcl6/MTA3/Mi-2/NuRD 复合物的抑制功能同时取决于 Bcl6 与被抑制的基因座上的组蛋白赖氨酸残基的乙酰化水平。Bcl6 的中心功能域需要被去乙酰化以促进它与 MTA3 的相互作用，而由 MTA3/Mi-2/NuRD 介导的基因抑制取决于Ⅰ型组蛋白去乙酰化酶 HDAC1 和 HDAC2（Fujita et al. 2004）。另外，Bcl6 通过它的 N 端 POZ 功能域与三个辅抑制子 SMRT、NCoR 和 BCoR 以一种排他性的方式相结合（Huynh and Bardwell 1998；Huynh et al. 2000）。这三种辅抑制子还能与Ⅱ型酶 HDAC3 相互作用，这可能能够增强 MTA3 介导的 Bcl6 靶基因的抑制或者使生发中心 B 细胞的另外一些基因发生沉默。

抗原对 B 细胞生发中心高亲和力 Ig 受体的刺激伴随着 Bcl6 蛋白表达水平的降低。受体的激活导致 MAPK 诱导的 Bcl6 磷酸化，使 Bcl6 蛋白通过泛素/蛋白酶体途径迅速被降解（Niu et al. 1998）。Bcl6 水平的下降解除了 *Prdm1* 基因的转录抑制，导致 Blimp1 蛋白的表达增加和随后的浆细胞发育（图 21-10）（Shaffer et al. 2000；Calame et al. 2003）。Blimp1 调控浆细胞分化的诸多方面。①Blimp1 通过抑制 *Pax5* 控制了 B 细

胞分化的核心转录程序（图 21-10）（Shaffer et al. 2002），Pax5 为维持 B 细胞功能和特征性所必需（Horcher et al. 2001；Mikkola et al. 2002）。②通过下调其他转录因子（如 Spi-B、EBF1、CIITA、Id3、Oct2 和 OBF1）的表达，Blimp1 间接终止了一些编码抗原受体信号通路和抗原呈递过程中的重要蛋白的基因表达（Shaffer et al. 2002）。③为了保证浆细胞的静息状态，Blimp1 直接抑制 *c-myc* 的转录（Shaffer et al. 2002）。④B 细胞中被 Pax5 抑制的世系不匹配基因，在浆细胞中因 Blimp1 介导的 Pax5 下调而被重新激活（Delogu et al. 2006）。因此，通过抑制其他抑制因子，Blimp1 可能间接诱导了其他维持浆细胞基本功能的基因的表达。⑤Blimp1 对于分泌型免疫球蛋白的表达是必须的（Calame et al. 2003），这些分泌型免疫球蛋白在内质网上的富集，作为未折叠蛋白应答通路的一部分，激活 XBP1 的表达。转录因子 XBP1 调控抗体的分泌，因此它对浆细胞的分化不可或缺（Reimold et al. 2001；Shaffer et al. 2004）。

有趣的是，Blimp1 也是早期胚胎发生过程中原始生殖细胞特化的关键决定因子，这在第 20 章已详细讨论。Blimp1 在发育的原始生殖细胞和浆细胞中介导转录抑制的机制尚不清楚。但是，在浆细胞中，Blimp1 有可能使用成纤维细胞中相同的抑制原理，在成纤维细胞中，小鼠 Blimp1 的人同源蛋白 PRD1-BF1 参与了病毒感染过程中干扰素 β 基因的诱导后抑制（Keller and Maniatis 1991）。多种机制参与了 PRDI-BF1 介导的 *IFNB1* 基因的抑制。PRD1-BF1 的结合能够置换掉 *IFNB1* 启动子上的转录激活因子（Keller and Maniatis 1991）。另外，PRD1-BF1 与 Groucho 蛋白家族辅抑制因子相互作用，利用组蛋白去乙酰化酶作为它们抑制机制的一部分（Ren et al. 1999）。PRD1-BF1 与 H3 第 9 位赖氨酸专一的组蛋白甲基转移酶亚家族成员 G9a 蛋白（Tachibana et al. 2002）的结合进一步导致沉默（Gyory et al. 2004）。Suv39H1 通过 H3K9me3 在着丝粒异染色质区域建立一个转录抑制的环境，与之不同的是，G9a 负责常染色质区域的 H3K9me2 和基因沉默（Tachibana et al. 2002）。G9a 的催化活性对于 PRDI-BF1 的抑制功能是必须的，因为无催化活性的 G9a 蛋白突变体失去了逆转 PRDI-BF1 对 *IFNB1* 进行转录抑制的能力（Gyory et al. 2004）。此外，G9a 与 PRDI-BF 相互作用区域的缺失会阻止 H3K9 的甲基化和 PRD-BF1 介导的转录沉默（Gyory et al. 2004）。鉴于以上这些信息，G9a 介导的组蛋白甲基化很可能是浆细胞中 Blimp1 建立稳定的基因表达模式的必需机制。有趣的是，Blimp1 蛋白也包含一个 PR（RIZ）亚家族的 SET 结构域。相关蛋白 RIZ1（Prdm2）的 SET 结构域被认为参与了肿瘤抑制和 H3 第 9 位赖氨酸的甲基化（Kim et al. 2003）。因此，Blimp1 的 SET 结构域在浆细胞分化过程中是否参与了基因抑制尚待揭晓。

## 4.2　成熟 B 细胞的发育可塑性

浆细胞的产生通常被认为是 B 细胞发育的终末阶段，因为免疫球蛋白基因的表达是 B 细胞和浆细胞的基本功能。有趣的是，免疫球蛋白基因的表达在这两种细胞类型中受通用转录因子而非 B 淋巴转录因子的协同控制。在免疫球蛋白基因之外，B 细胞和浆细胞则在基因表达模式上相距甚远（Shaffer et al. 2002，2004）。至于 Pax5 功能，在浆细胞中似乎截然相反，因为浆细胞特异性的 Pax5 沉默导致 B 细胞世系不匹配基因的重新激活，而这些基因在 B 细胞发育初期通常被 Pax5 所抑制（Delogu et al. 2006）。因

此，这些基因表达数据支持了一个可选择的观点，即抗原刺激的 B 细胞到浆细胞的分化是一个真正的世系转换。这一假说推断成熟 B 细胞的发育潜能应该是可塑的而并不是局限于 B 淋巴细胞的命运，以下证据支持了这一观点。

B 细胞转录因子 Bcl6 及其辅抑制因子 MTA3 在浆细胞细胞系中的人为表达抑制浆细胞特异性基因的表达，包括调节基因 *Blimp1* 和 *XBP1*（Fujita et al. 2004）。同时，多种 B 细胞特异性基因被激活，包括 Pax5 的靶基因 *CIITA* 和 *BLNK*，据推测也包括 *Pax5* 本身（Fujita et al. 2004）。因此，至少在体外培养条件下，Bcl6 和它的伙伴 MTA3，足以使浆细胞重启 B 细胞命运（图 21-11）。

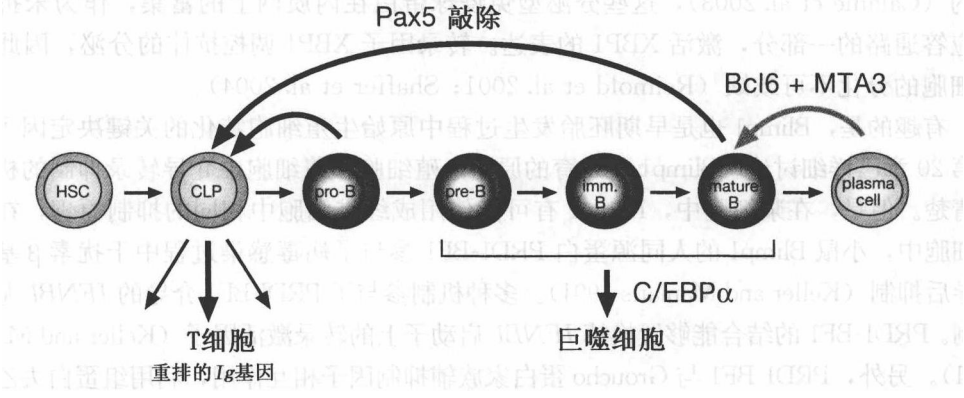



图 21-11 B 淋巴细胞的发育可塑性

人为表达 Bcl6 和 MTA3 在成熟的浆细胞细胞系中终止了浆细胞特异性基因的转录，同时重新激活了 B 细胞表达程序（桔色箭头）（Fujita et al. 2004）。$CD19^{+}$ B 淋巴细胞的发育阶段未被进一步研究，它能在强制表达 C/EBPα 的情况下在体外迅速转分化为巨噬细胞（红箭头）（Xie et al. 2004）。*Pax5* 的敲除使决定了的早 B 细胞甚至是成熟 B 细胞在体内反分化成未决定的源祖细胞，然后这些源祖细胞发育成骨髓中其他的造血细胞类型或是胸腺中的 T 细胞（黑箭头）（Mikkola et al. 2002；C. Cobaleda and M. Busslinger，未发表）。蓝色表示 B 细胞发育过程中 *Pax5* 的表达。

粒性白细胞发育的基本转录因子 C/EBPα 只在髓系始祖细胞及其分化后代细胞中表达（Akashi et al. 2000）。强行在骨髓或脾脏 B 细胞中表达 C/EBPα 导致被感染的 B 细胞在 5 天之内高效率地转分化为功能性巨噬细胞（Xie et al. 2004）。C/EBPα 从而激活了髓系基因程序，同时通过干扰 *Pax5* 的转录活性抑制 B 细胞特异性基因的表达（Xie et al. 2004）。因此，*Pax5* 功能的丧失可能作为对 C/EBPα 的人为表达的应答，而有助于 B 细胞向髓系世系的转变（图 21-11）。

条件性基因失活无可争辩地证明了 *Pax5* 在 B 淋巴细胞发生过程中控制 B 细胞特征性中的关键作用。在决定了的早 B 细胞中，Cre 介导的基因敲除证明 Pax5 不仅为 B 淋巴细胞转录程序的起始，并且为其在早期 B 细胞发育过程中的维持所必需（Mikkola et al. 2002）。作为 *Pax5* 失活的结果，初期决定了的早 B 细胞重新获得在体外分化成巨噬细胞和在体内发育成 T 细胞的能力（图 21-11）。成熟 B 细胞中 *Pax5* 的条件性缺失也导致 *Pax5* 依赖性基因表达程序的丢失（Horcher et al. 2001；Delogu et al. 2006）。但是更令人惊奇的是，将 *Pax5* 缺失的成熟 B 细胞注射到 RAG2 缺陷的小鼠体内后，它会

反分化成 *Pax5* 突变的源祖细胞。在 RAG2 缺陷的小鼠中，这些 *Pax5* 突变的源祖细胞回游到骨髓，再进入胸腺并完全重建 T 细胞的发育（图 21-11）。相应的 $CD4^+CD8^+$ 双阳性胸腺细胞中同时具有 *IgH*、*Igκ*、*TCRα* 和 *TCRβ* 重排的事实明确证明了 *Pax5* 缺失的成熟 B 细胞能够通过反分化到未决定的源祖细胞阶段从而转变成 T 细胞（C. Cobaleda and M. Busslinger，未发表*）。因此，*Pax5* 的表达对于从早 B 细胞到成熟 B 细胞阶段的发育过程中 B 淋巴细胞特征性的维持是必需的。通过对果蝇和脊椎动物发育系统的研究，人们认为转录因子能够通过改变基因表达模式来起始细胞命运决定，而决定了的细胞的转录状态是靠 *Polycomb* 和 *Trithorax* 基因编码的表观遗传因子来维持的（Ringrose and Paro 2004；第 11、12 章有更进一步的讨论）。对于转录因子 Pax5 的永久需求似乎驳斥了这些表观遗传记忆系统在 B 细胞发育过程中发挥的重要作用。然而，更为可能的是，Pax5 作为一个重要的募集因子，将 Polycomb 或 Trithorax 蛋白复合物募集到基因调节元件上从而维持 B 细胞的特征性。

## 5. 结　语

总的来说，多种表观遗传学机制参与了淋巴细胞发育的调节和导向作用。在所有不同的表观遗传调节因子中，我们目前了解最多的是转录因子的作用，它们通过将染色质修饰活性（如组蛋白乙酰转移酶或去乙酰化酶）募集到基因调控元件上从而控制整个基因表达模式。知道较少的是由组蛋白甲基转移酶、Trithorax 和 Polycomb 类蛋白或是 microRNA 和 siRNA 通路来控制的基因表达。分析这些调控系统的作用需要设计基因的条件敲除，因为组蛋白甲基转移酶、Trithorax 和 Polycomb 蛋白，还包括 RNAi 机器的组分不仅对淋巴细胞发生而且还对胚胎发育有非常重要的作用。此外，全基因组的 ChIP-on-chip 技术的发展和实用性使得不同淋巴细胞发生阶段整个染色体和复杂基因座（如抗原受体基因座）上的表观遗传修饰有可能被高分辨率地绘出。近来的这些进展为淋巴细胞发育的表观遗传学调控机制提供了重要的新观点。

吴　慧　译
张　慧　校
朱　冰　审校

## 参 考 文 献

Adolfsson J., Månsson R., Buza-Vidas N., Hultquist A., Liuba K., Jensen C.T., Bryder D., Yang L., Borge O.J., Thoren L.A.M., et al. 2005. Identification of Flt3$^+$ lympho-myeloid stem cells lacking erythro-megakaryocytic potential: A revised road map for adult blood lineage commitment. *Cell* **121:** 295–306.

Akashi K., Traver D., Miyamoto T., and Weissman I.L. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404:** 193–197.

Allman D., Sambandam A., Kim S., Miller J.P., Pagan A., Well D., Meraz A., and Bhandoola A. 2003. Thymopoiesis independent of common lymphoid progenitors. *Nature Immunol.* **4:** 168–174.

Barlev N.A., Emelyanov A.V., Castagnino P., Zegerman P., Bannister A.J., Sepulveda M.A., Robert F., Tora L., Kouzarides T., Birshtein B.K., and Berger S.L. 2003. A novel human Ada2 homologue func-

* 原书出版后，此工作被发表于 Nature 2007，449：473～477——译者

tions with Gcn5 or Brg1 to coactivate transcription. *Mol. Cell. Biol.* **23:** 6944–6957.

Bassing C.H., Swat W., and Alt F.W. 2002. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* (suppl.) **109:** S45–S55.

Baxter J., Merkenschlager M., and Fisher A.G. 2002. Nuclear organisation and gene expression. *Curr. Opin. Cell Biol.* **14:** 372–376.

Baxter J., Sauer S., Peters A., John R., Williams R., Caparros M.L., Arney K., Otte A., Jenuwein T., Merkenschlager M., and Fisher A.G. 2004. Histone hypomethylation is an indicator of epigenetic plasticity in quiescent lymphocytes. *EMBO J.* **23:** 4462–4472.

Bereshchenko O.R., Gu W., and Dalla-Favera R. 2002. Acetylation inactivates the transcriptional repressor BCL6. *Nat. Genet.* **32:** 606–613.

Bolland D.J., Wood A.L., Johnston C.M., Bunting S.F., Morgan G., Chakalova L., Fraser P.J. and Corcoran A.E. 2004. Antisense intergenic transcription in V(D)J recombination. *Nat. Immunol.* **5:** 630–637.

Brown K.E., Guest S.S., Smale S.T., Hahm K., Merkenschlager M., and Fisher A.G. 1997. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91:** 845–854.

Busslinger M. 2004. Transcriptional control of early B cell development. *Annu. Rev. Immunol.* **22:** 55–79.

Calame K.L., Lin K.I., and Tunyaplin C. 2003. Regulatory mechanisms that determine the development and function of plasma cells. *Annu. Rev. Immunol.* **21:** 205–230.

Chow C.M., Georgiou A., Szutorisz H., Maia e Silva A., Pombo A., Barahona I., Dargelos E., Canzonetta C., and Dillon N. 2005. Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. *EMBO Rep.* **6:** 354–360.

Chowdhury D. and Sen R. 2001. Stepwise activation of the immunoglobulin μ heavy chain gene locus. *EMBO J.* **20:** 6394–6403.

———. 2003. Transient IL-7/IL-7R signaling provides a mechanism for feedback inhibition of immunoglobulin heavy chain gene rearrangements. *Immunity* **18:** 229–241.

Dakic A., Metcalf D., Di Rago L., Mifsud S., Wu L., and Nutt S.L. 2005. PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis. *J. Exp. Med.* **201:** 1487–1502.

Delogu A., Schebesta A., Sun Q., Aschenbrenner K., Perlot T., and Busslinger M. 2006. Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* **24:** 269–281.

Fisher A.G. 2002. Cellular identity and lineage choice. *Nat. Rev. Immunol.* **2:** 977–982.

Francis N.J., Kingston R.E., and Woodcock C.L. 2004. Chromatin compaction by a polycomb group protein complex. *Science* **306:** 1574–1577.

Fujita N., Jaye D.L., Geigerman C., Akyildiz A., Mooney M.R., Boss J.M., and Wade P.A. 2004. MTA3 and the Mi-2/NuRD complex regulate cell fate during B lymphocyte differentiation. *Cell* **119:** 75–86.

Fukuda T., Yoshida T., Okada S., Hatano M., Miki T., Ishibashi K., Okabe S., Koseki H., Hirosawa S., Taniguchi M., et al. 1997. Disruption of the *Bcl6* gene results in an impaired germinal center formation. *J. Exp. Med.* **186:** 439–448.

Fuxa M., Skok J., Souabni A., Salvagiotto G., Roldán E., and Busslinger M. 2004. Pax5 induces *V-to-DJ* rearrangements and locus contraction of the *immunoglobulin heavy-chain* gene. *Genes Dev.* **18:** 411–422.

Goldmit M., Ji Y., Skok J., Roldan E., Jung S., Cedar H., and Bergman Y. 2005. Epigenetic ontogeny of the *Igk* locus during B cell development. *Nat. Immunol.* **6:** 198–203.

Goldmit M., Schlissel M., Cedar H., and Bergman Y. 2002. Differential accessibility at the κ chain locus plays a role in allelic exclusion. *EMBO J.* **21:** 5255–5261.

Grawunder U., Leu T.M.J., Schatz D.G., Werner A., Rolink A.G., Melchers F., and Winkler T.H. 1995. Down-regulation of *RAG1* and *RAG2* gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. *Immunity* **3:** 601–608.

György I., Wu J., Fejér G., Seto E., and Wright K.L. 2004. PRDI-BF1 recruits the histone H3 methyltransferase G9a in transcriptional silencing. *Nat. Immunol.* **5:** 299–308.

Hesslein D.G.T., Pflugh D.L., Chowdhury D., Bothwell A.L.M., Sen R., and Schatz D.G. 2003. *Pax5* is required for recombination of transcribed, acetylated, 5′ IgH V gene segments. *Genes Dev.* **17:** 37–42.

Honjo T., Muramatsu M., and Fagarasan S. 2004. AID: How does it aid antibody diversity? *Immunity* **20:** 659–668.

Horcher M., Souabni A., and Busslinger M. 2001. Pax5/BSAP maintains the identity of B cells in late B lymphopoiesis. *Immunity* **14:** 779–790.

Huynh K.D. and Bardwell V.J. 1998. The BCL-6 POZ domain and other POZ domains interact with the co-repressors N-CoR and SMRT. *Oncogene* **17:** 2473–2484.

Huynh K.D., Fischle W., Verdin E., and Bardwell V.J. 2000. BCoR, a novel corepressor involved in BCL-6 repression. *Genes Dev.* **14:** 1810–1823.

Ikawa T., Kawamoto H., Wright L.Y.T., and Murre C. 2004. Long-term cultured E2A-deficient hematopoietic progenitor cells are pluripotent. *Immunity* **20:** 349–360.

Jankovic M., Casellas R., Yannoutsos N., Wardemann H., and Nussenzweig M.C. 2004. RAGs and regulation of autoantibodies. *Annu. Rev. Immunol.* **22:** 485–501.

Keller A.D. and Maniatis T. 1991. Identification and characterization of a novel repressor of β-interferon gene expression. *Genes Dev.* **5:** 868–879.

Kim K.C., Geng L., and Huang S. 2003. Inactivation of a histone methyltransferase by mutations in human cancers. *Cancer Res.* **63:** 7619–7623.

Kitamura D. and Rajewsky K. 1992. Targeted disruption of μ chain membrane exon causes loss of heavy-chain allelic exclusion. *Nature* **356:** 154–156.

Kosak S.T., Skok J.A., Medina K.L., Riblet R., Le Beau M.M., Fisher A.G., and Singh H. 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296:** 158–162.

Kovanen P.E. and Leonard W.J. 2004. Cytokines and immunodeficiency diseases: Critical roles of the $\gamma_c$-dependent cytokines interleukins 2, 4, 7, 9, 15, and 21, and their signaling pathways. *Immunol. Rev.* **202:** 67–83.

Krangel M.S. 2003. Gene segment selection in V(D)J recombination: Accessibility and beyond. *Nat. Immunol.* **4:** 624–630.

Kraus M., Alimzhanov M.B., Rajewsky N., and Rajewsky K. 2004. Survival of resting mature B lymphocytes depends on BCR signaling via the Igα/β heterodimer. *Cell* **117:** 787–800.

Lee A.I., Fugmann S.D., Cowell L.G., Ptaszek L.M., Kelsoe G., and Schatz D.G. 2003. A functional analysis of the spacer of V(D)J recombination signal sequences. *PLoS Biol.* **1:** 56–69.

Liang H.E., Hsu L.Y., Cado D., and Schlissel M.S. 2004. Variegated transcriptional activation of the immunoglobulin κ locus in pre-B cells contributes to the allelic exclusion of light-chain expression. *Cell* **118:** 19–29.

Linderson Y., Eberhard D., Malin S., Johansson A., Busslinger M., and Pettersson S. 2004. Corecruitment of the Grg4 repressor by PU.1 is critical for Pax5-mediated repression of B-cell-specific genes. *EMBO Rep.* **5:** 291–296.

Maier H., Colbert J., Fitzsimmons D., Clark D.R., and Hagman J. 2003. Activation of the early B-cell-specific *mb-1* (Ig-α) gene by Pax-5 is dependent on an unmethylated Ets binding site. *Mol. Cell. Biol.* **23:** 1946–1960.

Maier H., Ostraat R., Gao H., Fields S., Shinton S.A., Medina K.L., Ikawa T., Murre C., Singh H., Hardy R.R., and Hagman J. 2004. Early B cell factor cooperates with Runx1 and mediates epigenetic changes associated with *mb-1* transcription. *Nat. Immunol.* **5:** 1069–1077.

Massari M.E., Grant P.A., Pray-Grant M.G., Berger S.L., Workman J.L., and Murre C. 1999. A conserved motif present in a class of helix-loop-helix proteins activates transcription by direct recruitment of the SAGA complex. *Mol. Cell.* **4:** 63–73.

McBlane F. and Boyes J. 2000. Stimulation of V(D)J recombination by histone acetylation. *Curr. Biol.* **10:** 483–486.

McMurry M.T. and Krangel M.S. 2000. A role for histone acetylation in the developmental regulation of V(D)J recombination. *Science* **287:** 495–498.

Medina K.L., Pongubala J.M., Reddy K.L., Lancki D.W., DeKoter R., Kieslinger M., Grosschedl R., and Singh H. 2004. Assembling a gene regulatory network for specification of the B cell fate. *Dev. Cell* **7:** 607–617.

Mikkola I., Heavey B., Horcher M., and Busslinger M. 2002. Reversion of B cell commitment upon loss of *Pax5* expression. *Science* **297:** 110–113.

Mito Y., Henikoff J.G., and Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat. Genet.* **37:** 1090–1097.

Miyamoto T., Iwasaki H., Reizis B., Ye M., Graf T., Weissman I.L., and Akashi K. 2002. Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment. *Dev. Cell* **3:** 137–147.

Morshead K.B., Ciccone D.N., Taverna S.D., Allis C.D., and Oettinger M.A. 2003. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. *Proc. Natl. Acad. Sci. USA* **100:** 11577–11582.

Mostoslavsky R., Singh N., Kirillov A., Pelanda R., Cedar H., Chess A., and Bergman Y. 1998. κ chain monoallelic demethylation and the establishment of allelic exclusion. *Genes Dev.* **12:** 1801–1811.

Mostoslavsky R., Singh N., Tenzen T., Goldmit M., Gabay C., Elizur S., Qi P., Reubinoff B.E., Chess A., Cedar H., and Bergman Y. 2001. Asynchronous replication and allelic exclusion in the immune system. *Nature* **414:** 221–225.

Niu H., Ye B.H., and Dalla-Favera R. 1998. Antigen receptor signaling induces MAP kinase-mediated phosphorylation and degradation of the BCL-6 transcription factor. *Genes Dev.* **12:** 1953–1961.

Nutt S.L., Heavey B., Rolink A.G., and Busslinger M. 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401:** 556–562.

Nutt S.L., Urbánek P., Rolink A., and Busslinger M. 1997. Essential functions of Pax5 (BSAP) in pro-B cell development: Difference between fetal and adult B lymphopoiesis and reduced *V*-to-*DJ* recombination at the *IgH* locus. *Genes Dev.* **11:** 476–491.

Orphanides G. and Reinberg D. 2000. RNA polymerase II elongation through chromatin. *Nature* **407:** 471–475.

Osipovich O., Milley R., Meade A., Tachibana M., Shinkai Y., Krangel M.S., and Oltz E.M. 2004. Targeted inhibition of V(D)J recombination by a histone methyltransferase. *Nat. Immunol.* **5:** 309–316.

Peters A.H., O'Carroll D., Scherthan H., Mechtler K., Sauer S., Schöfer C., Weipoltshammer K., Pagani M., Lachner M., Kohlmaier A., et al. 2001. Loss of the *Suv39h* histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107:** 323–337.

Rajewsky K. 1996. Clonal selection and learning in the antibody system. *Nature* **381:** 751–758.

Reimold A.M., Iwakoshi N.N., Manis J., Vallabhajosyula P., Szomolanyi-Tsuda E., Gravallese E.M., Friend D., Grusby M.J., Alt F., and Glimcher L.H. 2001. Plasma cell differentiation requires the transcription factor XBP-1. *Nature* **412:** 300–307.

Ren B., Chee K.J., Kim T.H., and Maniatis T. 1999. PRDI-BF1/Blimp-1 repression is mediated by corepressors of the Groucho family of proteins. *Genes Dev.* **13:** 125–137.

Ringrose L. and Paro R. 2004. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38:** 413–443.

Roldán E., Fuxa M., Chong W., Martinez D., Novatchkova M., Busslinger M., and Skok J.A. 2005. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat. Immunol.* **6:** 31–41.

Romanow W.J., Langerak A.W., Goebel P., Wolvers-Tettero I.L.M., van Dongen J.J.M., Feeney A.J., and Murre C. 2000. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. *Mol. Cell* **5:** 343–353.

Sato H., Saito-Ohara F., Inazawa J., and Kudo A. 2004. Pax-5 is essential for κ sterile transcription during Igκ chain gene rearrangement. *J. Immunol.* **172:** 4858–4865.

Schlissel M.S., Durum S.D., and Muegge K. 2000. The interleukin 7 receptor is required for T cell receptor γ locus accessibility to the V(D)J recombinase. *J. Exp. Med.* **191:** 1045–1050.

Seet C.S., Brumbaugh R.L., and Kee B.L. 2004. Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. *J. Exp. Med.* **199:** 1689–1700.

Shaffer A.L., Yu X., He Y., Boldrick J., Chan E.P., and Staudt L.M. 2000. BCL-6 represses genes that function in lymphocyte differentiation, inflammation and cell cycle control. *Immunity* **13:** 199–212.

Shaffer A.L., Lin K.I. Kuo T.C., Yu X., Hurt E.M., Rosenwald A., Giltnane J.M., Yang L., Zhao H., Calame K., and Staudt L.M. 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity* **17:** 51–62.

Shaffer A.L., Shapiro-Shelef M., Iwakoshi N.N., Lee A.-H., Qian S.B., Zhao H., Yu X., Yang L., Tan B.K., Rosenwald A., et al. 2004. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* **21:** 81–93.

Skok J.A., Brown K.E., Azuara V., Caparros M.L., Baxter J., Takacs K., Dillon N., Gray D., Perry R.P., Merkenschlager M., and Fisher A.G. 2001. Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. *Nat. Immunol.* **2:** 848–854.

Smale S.T. 2003. The establishment and maintenance of lymphocyte identity through gene silencing. *Nat. Immunol.* **4:** 607–615.

Stanhope-Baker P., Hudson K.M., Shaffer A.L., Constantinescu A., and Schlissel M.S. 1996. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. *Cell* **85:** 887–897.

Su I.H., Basavaraj A., Krutchinsky A.N., Hobert O., Ullrich A., Chait B.T., and Tarakhovsky A. 2003. Ezh2 controls B cell development through histone H3 methylation and *Igh* rearrangement. *Nat. Immunol.* **4:** 124–131.

Su I.H., Dobenecker M.W., Dickinson E., Osler M., Basavaraj A., Marqueron R., Viale A., Reinberg D., Wülfing C., and Tarakhovsky A. 2005. Polycomb group protein Ezh2 controls actin polymerization and cell signaling. *Cell* **121:** 425–436.

Tachibana M., Sugimoto K., Nozaki M., Ueda J., Ohta T., Ohki M., Fukuda M., Takeda N., Niida H., Kato H., and Shinkai Y. 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* **16:** 1779–1791.

Turner C.A.J., Mack D.H., and Davis M.M. 1994. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell* **77:** 297–306.

Vosshenrich C.A.J., Cumano A., Müller W., Di Santo J.P., and Vieira P. 2003. Thymic stroma-derived lymphopoietin distinguishes fetal from adult B cell development. *Nat. Immunol.* **4:** 773–779.

Waskow C., Paul S., Haller C., Gassmann M., and Rodewald H. 2002. Viable c-Kit$^{W/W}$ mutants reveal pivotal role for c-Kit in the maintenance of lymphopoiesis. *Immunity* **17:** 277–288.

West K.L., Singha N.C., De Ioannes P., Lacomis L., Erdjument-Bromage H., Tempst P., and Cortes P. 2005. A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. *Immunity* **23:** 203–212.

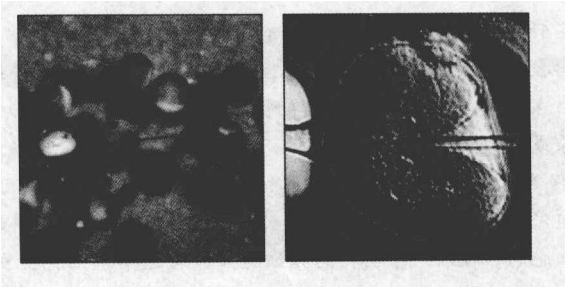
Xie H., Ye M., Feng R., and Graf T. 2004. Stepwise reprogramming of B cells into macrophages. *Cell* **117:** 663–676.

Yancopoulos G.D. and Alt F.W. 1985. Developmentally controlled and tissue-specific expression of unrearranged $V_H$ gene segments. *Cell* **40:** 271–281.

Ye B.H., Cattoretti G., Shen Q., Zhang J., Hawe N., de Waard R., Leung C., Nouri-Shirazi M., Orazi A., Chaganti R.S.K., et al. 1997. The *BCL-6* proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nature Genet.* **16:** 161–170.

Zhang S., Fukuda S., Lee Y., Hangoc G., Cooper S., Spolski R., Leonard W.J., and Broxmeyer H.E. 2000. Essential role of signal transducer and activator of transcription (Stat)5a but not Stat5b for Flt3-dependent signaling. *J. Exp. Med.* **192:** 719–728.

# 第 22 章



# 核移植和基因组重编程

Rudolf Jaenisch[1] and John Gurdon[2]

[1] *Whitehead Institute, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142-1479*

[2] *The Wellcome Trust /Cancer Research UK Gurdon Institute, The Henry Wellcome Building of Cancer and Developmental Biology, University of Cambridge, Cambridge CB2 1QN, United Kingdom*

## 本章目录

1. 历史
2. 核移植过程
   2.1 两栖类动物
   2.2 哺乳动物
3. 克隆动物的表型
   3.1 两栖动物
   3.2 哺乳动物
   3.3 从终末分化细胞中获得克隆的哺乳动物
4. 核重编程伴随的变化
   4.1 两栖类
   4.2 哺乳动物
5. 表观遗传记忆
6. 核移植的医学应用
   6.1 繁殖性克隆
   6.2 核移植的治疗性应用
   6.3 繁殖性克隆与治疗性克隆的区别

参考文献

## 概要

动物个体由行使多种生理功能的上百种不同类型的细胞构建而成。一个很早就提出的关键问题是如何确保基因差异性表达，即保证正确的基因活化或沉默，从而使得给定的分化细胞能够维持其正常功能。早先，在基因表达的分子基础得到重视之前，曾有假设认为组织特异性基因表达的基础可能是遗传性去除或永久失活那些组织里的沉默基因。确实，在某些生物体中，如昆虫 *Sciara*，遗传物质只完整保存于生殖系细胞中而在体细胞中被去除。这就提出了“核均一性”的问题，即是否体细胞的基因组保存了完整的遗传信息。针对这个问题最普遍的解决途径是核移植，也就是通过将一个供体体细胞的核植入一个去核的卵细胞来检验这个体细胞核发育成新生物体的潜能。的确，体细胞核来源的克隆动物的成功无疑证明了阻止体细胞核生成所有类型细胞的重大遗传变化并不在正常发育过程中发生。

## 1. 历　　史

在多细胞生物中，首次将一个细胞的细胞核成功的移植入另一个细胞是由 Briggs 和 King 在 1952 年完成（Briggs and King 1952），虽然此前核移植已经在包括阿米巴虫、纤毛虫和 Acetabularia（Gurdon 1964）内的单细胞生物体中获得成功。Briggs 和 King（1952）将一个豹蛙（*Rana pipiens*）的囊胚卵泡细胞核移植入一个去核的卵细胞，并获得了正常游动的蝌蚪。在 *Rana pipiens* 中的该研究以及后续工作中，约 30% 的囊胚核移植发育到形态正常的后神经轴胚阶段。这一早期成功的重要性在于它为检测分化细胞的细胞核是否也能支持受体卵的正常发育开辟了道路，即替换正常受精卵的核。由 Briggs 和 King 发表的另一篇重要文献（1957）报道了在囊胚阶段之后非常短的时间内，体细胞（在本例中为内胚层）的核失去了支持正常发育的能力。确实，到尾-芽阶段，内胚层的核不再能支持任何正常的发育。

此后不久，在爪蟾中报道了成功的核移植（Fischberg et al. 1958）。在爪蟾中，一个遗传标记被用来证明核移植产生的胚胎完全来源于植入核的活性而非被紫外照射致死的受体卵的核。核移植胚胎发育在爪蟾中比豹蛙中更正常，并且很快发育到成体阶段。通过在爪蟾中移植入内胚层细胞的核获得了首批性成熟的成体克隆动物（Gurdon et al. 1958），并且绝大多数在各方面均表现正常。

核移植要想获得成功，一个卵细胞必须“重编程”体细胞供体核到一个“胚胎的”表观遗传阶段从而使得从胚胎发育所必需的遗传程序可以被活化。这是现阶段研究中为了解核移植实验中重编程分子机制的焦点问题。由于两栖类和哺乳类的胚胎发育非常不同，各系统可以为核移植实验中重编程问题提供不同的线索。例如，从两栖类动物中可以获得大量的体积很大的卵，该系统在生化研究中尤其有用。相对应的，哺乳动物，尤其是小鼠，允许进行组织培养和遗传研究。因为这些原因，以下将在两栖类和哺乳类系统中的观察互为补充，且并列讨论以强调其相似性和不同点。首先，我们描述克隆所得的两栖类和哺乳类的表型以及其如何被供体核的不同状态所影响。随后讨论被认为对重编程有重要作用的分子机制。最后讨论治疗性核移植技术的潜在应用。

## 2. 核移植过程

### 2.1　两栖类动物

爪蟾卵的全年可获得性以及水生动物在实验室的易维持性，使得两栖动物核移植工作在豹蛙中首次成功之后，更多地在爪蟾中进行。根据核注射入去核的卵或者未去核的卵母细胞，可以将实验分成两类（图 22-1）。核移植需要供体细胞的质膜经处理具有通透性，这种处理或者是通过机械手段，将细胞吸入比细胞小的吸管；或者是通过化学手段，将细胞短暂的暴露于一个能与膜整合的药物。供体细胞随核引入的细胞质含量仅约为卵体积的 $10^{-5}$，几乎没有任何作用。移植后的卵细胞培养于简单的无营养的盐溶液

中，因而其发育不依赖于培养液。

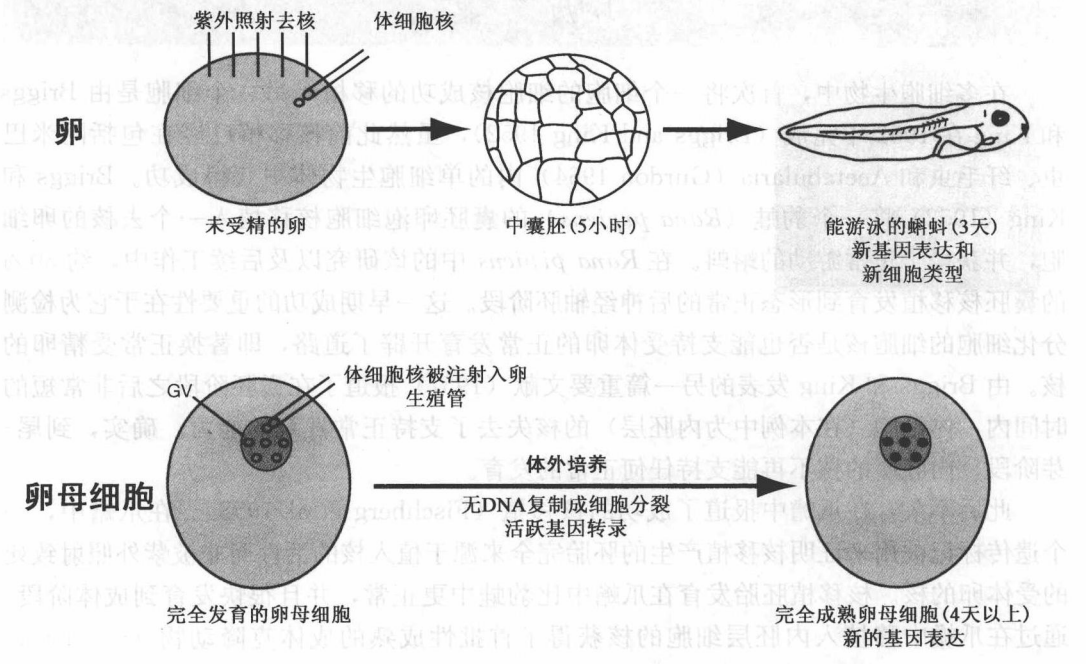



图 22-1 爪蟾中两类核移植实验示意图

上图显示了单个核植入紫外（UV）照射去核的卵的过程。核移植 3 天后形成可游动的蝌蚪。下图显示了将多个核植入处于第一次减数分裂前期生长中的卵母细胞胚泡或卵的核（GV）的过程。被注入核的卵并不分裂，并且可以培养许多天。被注入的体细胞核在培养的开始 4 天在基因表达上发生变化。

核移植入卵母细胞则完全不同。从雌性卵巢中取出的完全发育的卵母细胞处于在第一次减数分裂的前期（图 22-1）。约 50 个体细胞核被注射入这些发育中的卵母细胞的细胞核（生殖管），它们不复制 DNA 或也不分裂，而在几天时间内在转录上变得逐渐的活跃。获得移植核的卵母细胞像上述的卵一样，在无营养的盐溶液中培养，并且在 2 周或更长时间的整个培养过程中不发生形态上的变化。

## 2.2 哺乳动物

最早的哺乳动物克隆尝试是在家兔中进行的。在这些实验中，卵母细胞与桑椹胚时期的胚胎细胞融合，得到的三倍体克隆可以发生几次卵裂（Bromhall 1975）。在 1984 年，McGrath 和 Solter 通过 Sendai 病毒介导的细胞融合，有效地将供体细胞核移入去核的受精卵中，该方法被用于以下两类实验：①获取单亲胚胎，以雄原核替换雌原核得到单亲雄胚胎，或是以雌原核替换雄原核得到单亲雌胚胎（McGrath and Solter 1984a，b；Surani et al. 1984）。单性胚胎毫无例外的不能发育，这也是第一次证明了在正常胚胎发育过程中，亲本特异的基因组印记的重要作用。②将卵裂期胚胎的供体核移植入去核的受精卵中获得克隆胚胎。没有一个重建的胚胎发育超过卵裂后期，这导致一个结论认为基因组的全能性在早期发育过程中被很快地丢失，与两栖动物相比，几乎预示了不

可能得到哺乳动物的克隆（McGrath and Solter 1984b）。但是，从其他哺乳动物中得到的实验结果很快就挑战了这一结论。

1986 年，Willadsen 用 4～16 细胞时期的供体胚胎核进行核移植成功得到了活的克隆羊，并且用这种方法很快得到了克隆的牛和猪（Robi et al. 1987；Prather et al. 1989）。为何是农场动物克隆而非对照严格的小鼠克隆实验得到了成功呢（McGrath and Solter 1984b）？这些物种间发育上一个主要不同是从母源发育控制（依赖于母源储存的 RNA）转变为合子发育控制（依赖于合子产生的 RNA）的时机。在羊和牛的胚胎中这个主要的转变看起来发生于 8～16 细胞时期（Calarco and McLren 1976；Camous et al. 1986），但是在小鼠胚胎中在 2 细胞阶段已经完成了（Bolton et al. 1984）。因此，在克隆羊或牛胚胎中比在克隆小鼠胚胎中，激活供体基因组的时间限制更为宽松。而在小鼠中，供体核必须在核移植后被迅速激活才能使卵裂得以持续。

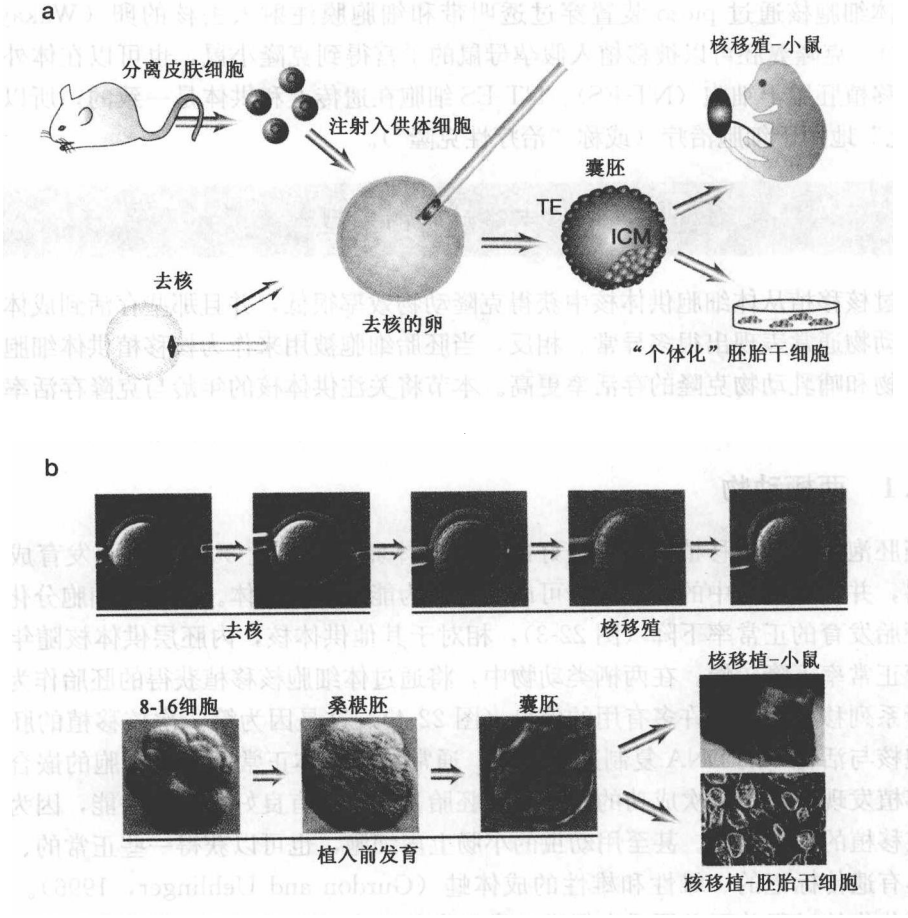



图 22-2　通过显微注射进行小鼠核移植

（a）核移植过程概要图。当把囊胚种于照射后的饲养层细胞上时，内细胞团（ICM）可生成 ES 细胞；当植入同步的假孕母鼠中，囊胚可发育成为小鼠。囊胚的外细胞团，即滋养外胚层（TE）可发育成为胚外组织（胎盘），而 ICM 细胞发育成为胚胎。（b）光学显微镜观察到的 a 中的各个步骤（经许可重印，来自 Meissner 2006）。

由体细胞核移植而非从卵裂期胚胎中获得供体核进行核移植的首次成功是从体外培养的成纤维供体细胞获得克隆羊的实验（Campbell et al. 1996b）。紧接着是以乳腺细胞核为供体的克隆羊“多莉”的诞生（Wilmut et al. 1997），这是第一次从成年供体细胞获得的哺乳动物克隆。从此，共有 15 种哺乳动物，包括小鼠、山羊、猪、牛、兔子、大鼠、猫和狗等相继被克隆（见 Campbell et al. 2005 综述）。

与两栖动物克隆一样，哺乳动物的克隆都要经过两个步骤。绝大多数情况下，成功的体细胞核移植实验均是供体细胞核移植入去核的卵母细胞（与早期的小鼠克隆被移植入去核的受精卵不同；McGrath and Solter 1984）。首先，在减数分裂后期用吸管将接受移植的卵细胞的纺锤丝除去，同时将供体细胞核移植进核卵细胞。在大多数哺乳动物中，供体细胞核通过电融合的方式进入去核的卵细胞。但由于小鼠的卵细胞易碎，所以在核移植过程中，将供体核注射进入更为有效（Wakayama et al. 1998）。如图 22-2 所示，供体细胞核通过 piezo 装置穿过透明带和细胞膜注射入去核的卵（Wakayama et al. 1998）。克隆囊胚可以被移植入假孕母鼠的子宫得到克隆小鼠，也可以在体外组织培养成核移植胚胎干细胞（NT-ES）。NT-ES 细胞在遗传上和供体是一致的，所以，可以“个性化”地用于细胞治疗（或称“治疗性克隆”）。

## 3. 克隆动物的表型

通过核移植从体细胞供体核中获得克隆动物效率很低，并且那些存活到成体的稀少的克隆动物通常表现出很多异常。相反，当胚胎细胞被用来作为核移植供体细胞时，两栖类动物和哺乳动物克隆的存活率更高。本节将关注供体核的年龄与克隆存活率之间的相互关系以及从成体和胚胎供体细胞中获得的两栖类和哺乳动克隆的表型。

### 3.1 两栖动物

当胚泡细胞核被移植入质量良好的爪蟾受体卵时，超过 30%的胚胎发育成为正常的蝌蚪，并且它们之中的绝大部分可以发育成为能繁殖的成体。当供体细胞分化后，核移植胚胎发育的正常率下降（图 22-3），相对于其他供体核，内胚层供体核随年龄增加时发育正常率下降较慢。在两栖类动物中，将通过体细胞核移植获得的胚胎作为供体核再进行系列移植得到了许多有用的信息（图 22-4）。这是因为第一次核移植的胚胎由于体细胞核与活化卵的 DNA 复制速度不同，通常是染色体正常及异常细胞的嵌合体；系列核移植发现经历了一次成功的核移植的胚胎，其核具有良好的发育潜能，因为这些细胞受核移植的损伤最低。甚至用幼虫的小肠上皮细胞，也可以获得一些正常的、性成熟的、具有遗传标记的、雌性和雄性的成体蛙（Gurdon and Uehlinger，1996）。这一结果说明分化的过程并不必须丢失促进正常发育的能力，因而原则上阐明了细胞分化时仍可以保持基因组的完整性。

在两栖类核移植实验中，从一个成体核获得正常的成体克隆动物是不可能的。但是，可以从许多不同成体组织的细胞核得到形态正常的蝌蚪（Laskey and Gurdon 1970），并且这些蝌蚪含有正常功能范围的分化的细胞类型。因此，那些已经决定了分

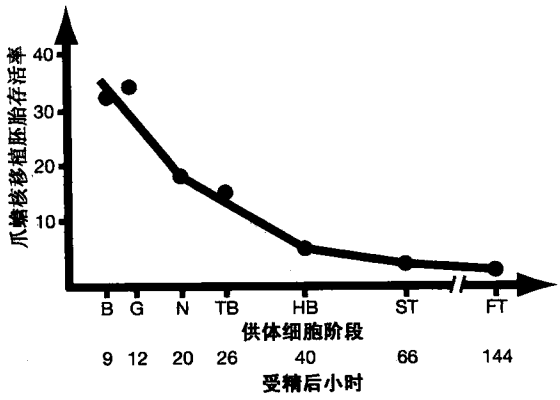



图 22-3　爪蟾核移植胚胎的存活率随着供体细胞分化程度的增加而降低

供体细胞所处时期缩写：（B）囊胚；（G）原肠胚；（N）神经轴胚；（TB）尾端；（HB）心脏；（ST）游动的蝌蚪；（FT）喂养期的蝌蚪（经许可重印自 Gurdon 1960）。

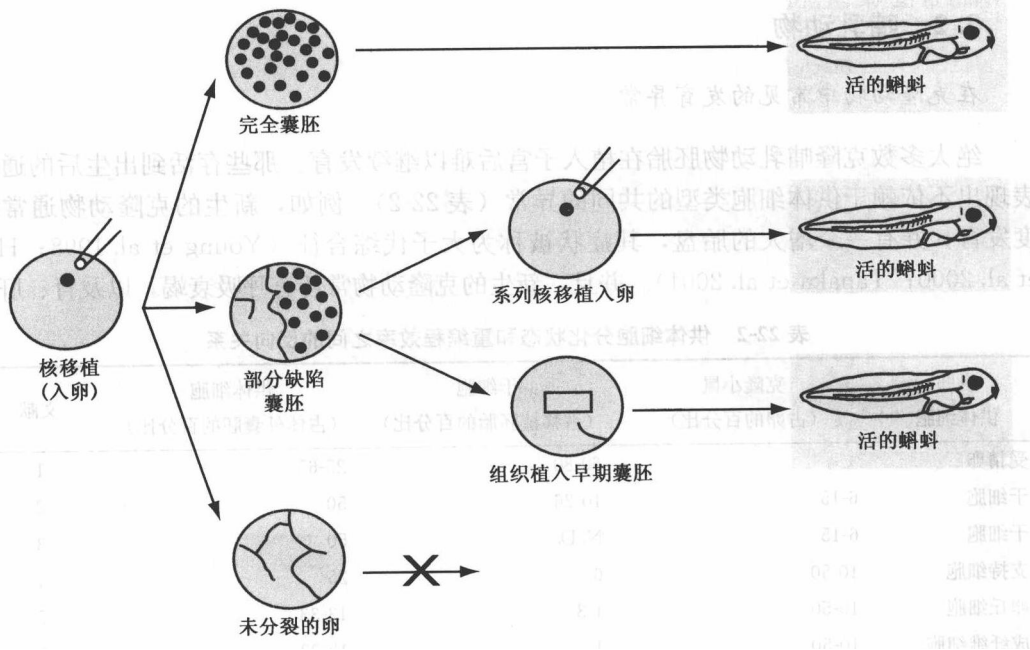



图 22-4　爪蟾中的系列核移植和嫁接

对于标准的第一次移植，单个体细胞核被注入一个去核的卵中，该胚胎直接发育成为蝌蚪。对于系列核移植，在发育早期的第一次移植的胚胎被用来为下一次核移植提供供体细胞核。这一系列核移植胚胎可长到蝌蚪阶段。最后，一个初次核移植的部分缺陷的胚胎被用来为移植入来自受精卵发育的宿主胚胎提供一片组织。该移植物可发育成蝌蚪的部分器官。

化方向的细胞和卵细胞质一起，包含了促进向绝大多数非相关细胞分化的遗传潜能。

从幼虫小肠内皮细胞核移植获得的胚胎可以对一种类型细胞核促进向其他类型细胞分化的能力进行定量，定量依据为判断其分化成具有功能的肌肉和神经的程度，而后者

可以通过观察胚胎受刺激后的游动来测定。这种胚胎通常可以通过对第一次核移植后得到的不完全正常的胚胎进行序列核移植而获得（图 22-4）（Gurdon1962）。并且，将从形态缺陷的初次核移植胚胎获得的胚泡细胞植入正常的从受精卵而来的受体时，该细胞可以发育成肌肉（图 22-4）（Byrne et al. 2003）。这个结果说明多达 30%的小肠内皮细胞可以在核移植后发育成为有功能的轴向肌肉细胞（表 22-1）。从两栖类核移植得到的发育异常类型与供体细胞来源形态水平间不呈任何相关性。无论供体核的细胞类型和发育阶段，核移植胚胎死亡于一个相似范围的缺陷，包括不完全卵裂，不能出腔，轴线形成的缺陷，以及缺乏头部结构。这些看起来杂乱的缺陷并不令人惊讶，因为初次核移植得到的胚胎细胞中常见大的染色体异常（见 4.1 节，克隆胚胎的重编程）。

**表 22-1 细胞核重编程效率:幼体爪蟾内胚层细胞**

| | | |
|---|---|---|
| 第一次核移植 | 胚胎的肌肉和神经 | 15% |
| 第一次和连续核移植 | 胚胎的肌肉和神经 | 22% |
| 第一次和连续核移植并嫁接于宿主 | 胚胎的肌肉和神经 | 30% |

## 3.2 哺乳动物

*在克隆动物中常见的发育异常*

绝大多数克隆哺乳动物胚胎在植入子宫后难以继续发育。那些存活到出生后的通常表现出不依赖于供体细胞类型的共同的异常（表 22-2）。例如，新生的克隆动物通常过度发育，并有一个增大的胎盘，其症状被称为大子代综合征（Young et al. 1998；Hill et al. 2000；Tanaka et al. 2001）。并且，新生的克隆动物常伴有呼吸衰竭，以及肾、肝、

**表 22-2 供体细胞分化状态和重编程效率之间的反向关系**

| 囊胚<br>供体细胞 | 克隆小鼠<br>（占卵的百分比） | 干细胞<br>（占移植胚胎的百分比） | 供体细胞<br>（占体外囊胚的百分比） | 文献 |
|---|---|---|---|---|
| 受精卵 | | 60-80 | 25-65 | 1 |
| 干细胞 | 6-15 | 10-26 | 50 | 2 |
| 干细胞 | 6-15 | N. D. | 50 | 3 |
| 支持细胞 | 10-50 | 6 | 25 | 4 |
| 卵丘细胞 | 10-50 | 1-3 | 13-33 | 5 |
| 成纤维细胞 | 10-50 | 1 | 13-33 | 5 |
| NKT | 70 | 4 | N. D. | 6 |
| B,T 细胞 | 4 | N. D. | 7 | 7 |
| 神经元 | 15 | N. D. | 6-28 | 8 |
| 黑素瘤细胞 | 1.5 | N. D. | 25 | 9 |

注:N. D 表示未确定。

引用：[1]Wakayama et al. 2005. [2]Wakayama et al. 1999；Rideout et al. 2000；Eggan et al. 2001；Humpherys et al. 2001；[3]Blelloch et al. 2004. [4]Ogura et al. 2000. [5]Wakayama et al. 1998，2005；Wakayama and Yanagimachi 1999. [6]Inoue et al. 2005. [7]Hochedlinger and Jaenisch 2002. [8]Eggan et al. 2004；Li et al. 2004. [9]Hochedlinger et al. 2004.

心和脑的疾病。即便长期存活的个体也可能在生命后期出现异常。例如，老年克隆小鼠通常变得肥胖，出现严重的免疫疾病，或过早死亡（Ogonuki et al. 2002；Tamashiro et al. 2002）。如图 22-5 中表示的那样，绝大多数克隆无法通过刚植入子宫后和出生这两个关键期。这可能是对基因表达错误最为敏感的两个关键性发育时期。但是，成体动物的获得以及看起来健康的成体克隆动物已被作为核移植可以产生正常克隆动物的证据，即使其效率很低。重要的是，克隆动物中严重的异常经常只在动物衰老时才变得明显（Ogonuki et al. 2002；Tamashiro et al. 2002）。在许多甚至绝大部分成体克隆动物中随机发生的疾病和其他缺陷意味着克隆动物的存活并不能保证其“正常性”。相反，存活的克隆动物的表型在一个广泛的范围内分布，包括在出生后早期由异常导致的突然死亡或更多的微小异常使其得以存活到更高的年龄（图 22-5）。这些考虑说明确定细微的基因表达缺陷的复杂性，并强调了对更多精密的测试标准如环境压力和行为学测试的需求。

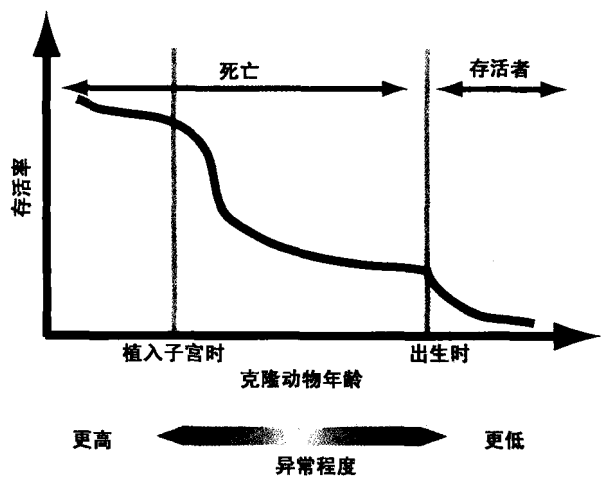



图 22-5　哺乳动物克隆的存活率

克隆动物表型的异常很广泛。大多数克隆发育阻滞在两个明显的阶段上——植入子宫时和出生时。更多基因的细小的表达差别导致发育后期的疾病和死亡。

*表观遗传原因与遗传原因*

克隆动物特有的异常并不被其后代所继承，说明“表观遗传”而非遗传的畸变是其原因。这是因为表观遗传变化，而非遗传变化，是在基因组从生殖系中传递时可被抹去的可逆的 DNA 或染色质修饰（Ogonuki et al. 2002；Tamashiro et al. 2002）。因此，克隆存在的问题是因为植入的供体核发生错误的“表观遗传/基因组重编程”，而非在供体核细胞中获得的体细胞突变。

## 3.3　从终末分化细胞中获得克隆的哺乳动物

*供体细胞分化状态和核重编程效率*

在两栖类核移植实验中已经提出的问题说明了在供体细胞核分化状态和其在植入卵

细胞中后发育的能力之间存在负相关性（图 22-3）。一个重要的问题是在哺乳动物中，供体细胞分化的状态是否也会影响重编程的效率。如表 22-2 中总结的那样，重编程可以在功能上通过评价克隆在几个不同时期水平的发育而测量，包括①在核移植入卵细胞后的囊胚形成率，②克隆胚胎在植入子宫后存活到出生或成年的比例，③从克隆的囊胚通过体外培养获得多能胚胎干细胞（ES）的比例。

重组胚胎到囊胚的植入前发育效率部分上受实验参数如细胞周期和被移植的细胞核的物理状态的影响。例如，非分裂阶段的供体细胞的克隆比活跃增殖期的细胞的克隆更有效率（Campbell et al. 1996a；Cibelli et al. 1998）。由此可以预期，由处于 $G_1$ 或 $G_0$ 细胞周期的成纤维细胞、睾丸支持细胞、卵丘细胞或 NKT 细胞重组得到的卵具有较高的发育到囊胚的比例，相反，对于大部分细胞处于 S 期的活跃分裂的 ES 或胚胎肿瘤（EC）细胞，其发育比例较低（表 22-2）。由于实验中的分裂期波动性较大，衡量由重组的卵发育而来的囊胚的比例并不是一个定量“重编程能力”的可信标准。但是，一旦一个克隆胚胎发育到了囊胚阶段，在植入子宫后到出生阶段的发育依赖于供体核的分化状态。由胚胎类供体如 ES 或 EC 细胞得到的克隆胚胎的发育效率比从卵丘或成纤维供体细胞来源的克隆发育效率高约 10～20 倍（Eggan et al. 2001；Wakayama and Yanagimachi 2001），想必是因为一个未分化的胚胎细胞的核比一个分化的体细胞核更易于，或更少需要重编程。这说明了在供体细胞分化程度和重编程效率之间存在负相关性。最后，一旦一个胚胎发育到囊胚阶段，其有相当一致的可能性来获得 ES 细胞，说明从外植的囊胚获得的 ES 细胞更少的依赖于供体核的分化状态（表 22-2）。

### 终末分化细胞的核可以被重编程到全能的状态么？

在早期的两栖类和哺乳类克隆实验中，用于核移植的供体核的种类是多样的，并且不能排除是存在于供体群中极少的成体干细胞而非分化细胞的核产生了少数存活的克隆。例如，成体干细胞的表观遗传状态可能与胚胎干细胞相似，并且可能更易于重编程因而可能更易于生成存活的克隆。为了解决是否终末分化的细胞核足以被重编程以产生成体动物的这个问题，需要遗传标记来确定一个成活的克隆供体核。这些标记被用来明确地证明从成熟的免疫细胞、从终末分化的神经元细胞和从恶性肿瘤细胞获得的核都可以被重编程并产生成体的克隆动物。

### 由成熟免疫细胞获得的单克隆小鼠

单克隆小鼠是通过外周淋巴细胞核移植产生的。在外周淋巴细胞中，免疫球蛋白（Ig）和 T 细胞受体（TCR）基因的遗传重排可以被用来作为表现某给定克隆供体核同一性和分化状态的稳定标记。因为早期试图产生单克隆小鼠的尝试是失败的，所以采用了两步克隆法，先从克隆囊胚中获得 ES 细胞，然后，用克隆的 ES 细胞获得单克隆小鼠（图 22-6）。从 B 或 T 细胞供体核获得的动物可以存活并且在所有组织中均携带完全重排的免疫球蛋白或 TCR 基因（Hochedlinger and Jaenisch 2002）。像预期的那样，单克隆小鼠的免疫细胞只表达那些在相应的被用作核移植的供体细胞中有效重排的 Ig 和 TCR 基因座的基因，而其他 Ig 和 TCR 基因的重排被抑制。这些结果明确地说明终末

分化供体细胞的核可以通过核移植被重编程到多能状态。直接从成熟的 B 或 T 细胞获得克隆胚胎（而非在我们实验中采取的两步法）的效率，虽然难以确定，但明显比由从成纤维细胞或卵丘细胞获得克隆的效率低（在 2000 个操作的胚胎中可能少于 1 个）（表 22-2）。近期，终末分化的 NKT 细胞已被直接克隆（Inoue et al. 2005）。NKT 细胞与 B 和 T 细胞类似，具有基因重排，从而可以反馈性地确定细胞的分化状态。但是，虽然 T 细胞和 NKT 细胞属于相同的细胞世系，他们相应的核移植效率差异很大（Hochedlinger and Jaenisch 2002；Inoue et al. 2005）。

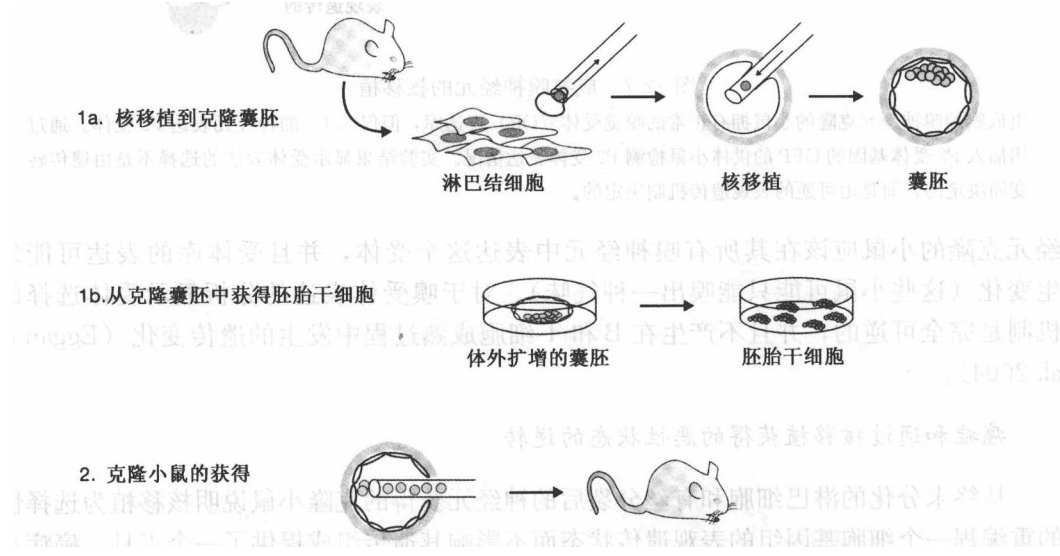



图 22-6　由成熟淋巴供体细胞获得单克隆小鼠的两步法

(1) 来源于外周淋巴结细胞的核被注入去核的卵，并得到克隆囊胚。囊胚经体外培养获得克隆的 ES 细胞。(2) 在第二步中，单克隆小鼠经四倍体互补实验获得（Eggan et al. 2001；Hochedlinger and Jaenisch 2002）。

### 由成熟的嗅神经元获得的克隆小鼠

与 B 和 T 细胞不同，有丝分裂后的神经元细胞核在其分化的程序中已经不可逆地退出了细胞周期。为了验证一个成熟神经元的细胞核是否可以被重编程到全能状态，通过采用和生成单克隆小鼠相似的方式，从嗅神经元中克隆了可育的成年小鼠（图 22-6）（Eggan et al. 2004；Li et al. 2004）。如表 22-2 中总结的那样，从嗅神经元细胞克隆 ES 细胞的效率与从免疫细胞核获得 ES 的效率在相同范围内。这些观察说明一个有丝分裂后的神经细胞核可以重新进入细胞周期并能被重编程到多能状态（图 22-7）。

在小鼠嗅上皮的 $2\times10^6$ 个细胞中，每个细胞只表达 1500 个嗅受体（OR）基因中的一个，因而一个神经元的功能特征性由其表达的受体的本质决定（这与在 21 章讨论到的 B 或 T 细胞中的免疫球蛋白或 TCR 基因单等位基因的表达相似）。允许单个嗅觉受体随机选择的一个机制可能是 DNA 重排。从一个成熟的嗅神经元克隆得到的小鼠使得研究嗅觉受体选择中是否发生了不可逆的 DNA 重排成为可能。如果嗅觉受体选择中发生了 DNA 重排，那么与上述描述的单克隆小鼠相似，由表达 P2 的神

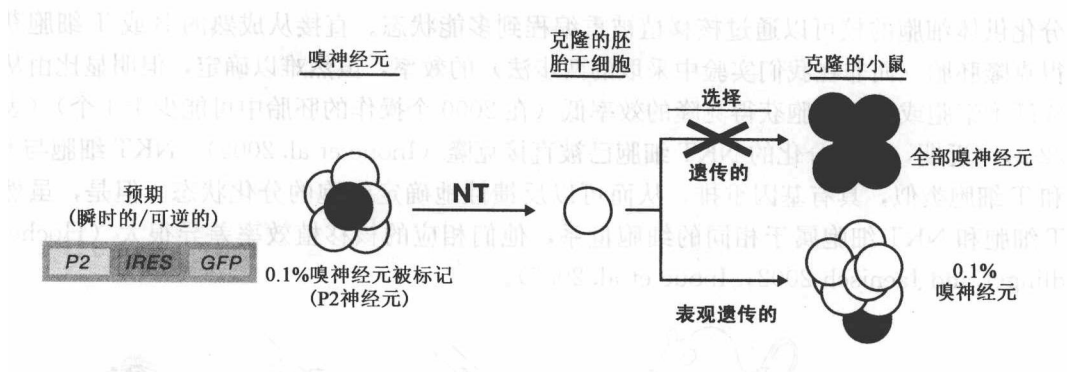



图 22-7 成熟嗅神经元的核移植

由成熟的嗅神经元克隆的小鼠拥有正常的嗅觉受体（OR）表达库，但仅 0.1%的神经元表达 P2 受体。通过用插入 P2 受体基因的 GFP 的供体小鼠检测 P2 受体表达情况。实验结果显示受体表达的选择不是由遗传转变而决定的，而是由可逆的表观遗传机制决定的。

经元克隆的小鼠应该在其所有嗅神经元中表达这个受体，并且受体库的表达可能发生变化（这些小鼠可能只能嗅出一种气味）。对于嗅受体表达的分析显示受体选择的机制是完全可逆的，并且不产生在 B 和 T 细胞成熟过程中发生的遗传变化（Eggan et al. 2004）。

### *癌症和通过核移植获得的恶性状态的逆转*

从终末分化的淋巴细胞和有丝分裂后的神经元获得的克隆小鼠说明核移植为选择性的重编程一个细胞基因组的表观遗传状态而不影响其遗传组成提供了一个工具。癌症是由遗传和表观遗传转变共同导致的，但是表观遗传在一个癌症细胞的恶性表型上的影响并不清楚。癌症供体细胞的核移植可用来作为工具检验转化状态的可逆性。确实，两栖类早期的实验证明来源于肾肿瘤细胞的核可以被重编程从而支持早期发育直到蝌蚪阶段（Mckinnell 1962）。相似的实验结果也在小鼠中获得。通过对髓母细胞瘤（medulloblastoma）细胞系的核移植获得的胚胎，虽然效率低，但仍可支持早期发育，直到发育阻滞（ Li et al. 2003）。但是，这些实验不能明确地说明由肿瘤细胞核移植得到的克隆不含有非转化细胞的污染。当多种肿瘤细胞的核，包括白血病、淋巴瘤、乳腺癌和黑色素瘤细胞核被移植到去核的小鼠卵细胞中，绝大多数可以支持到植入前发育，形成外观正常的囊胚（Hochedlinger et al. 2004）。因此，卵环境可以抑制这些肿瘤细胞的恶性表型，并支持看起来正常的早期发育。但是，只有一个来自 RAS 诱导的黑色素瘤模型的基因组可以形成能分化成嵌合小鼠中绝大多数体细胞世系的 ES 细胞系。但是，因为在供体细胞中存在遗传改变，所有的嵌合体都产生了癌症。这些发现说明癌细胞核在暴露于卵细胞质后可发育分化成所有细胞类型，说明癌症的恶性表型大部分是由表观遗传转变所决定。而从 EC 细胞供体核的克隆实验中得到了一个不同的结论：与体癌症细胞核不同，胚胎类肿瘤的恶性表型是由遗传转变导致的，因为其不能通过暴露于卵细胞质实现逆转（Blelloch et al. 2004）。

## 4. 核重编程伴随的变化

早期发育所采取的策略在两栖类和哺乳类中有明显不同。例如，爪蟾胚胎的卵裂是迅速的，每个细胞周期约为 30 分钟，哺乳动物胚胎正相反，其在受精后 24 小时内只分裂一次。并且，爪蟾的合子基因组只在 12 次有丝分裂周期后才在囊胚转变中期开始表达，而小鼠胚胎基因组在 2 细胞时期开始活化。因此，在两栖类和哺乳类中采取的不同发育策略会影响体细胞供体核的重编程就不足为怪了。本节对比了在爪蟾和哺乳动物克隆正常发育中发生的表观遗传重编程。一个有趣的问题是体细胞供体核的表观遗传状态是否影响克隆胚胎的基因表达模式，这被称为“表观遗传记忆”，本章将会继续讨论。

### 4.1　两栖类

在两栖类中，卵子和受精卵的核及染色体所处的状态与体细胞中完全不同。卵的胚泡包含有最大程度伸展的在转录水平上极度活化的灯刷状染色体（Callan and Lloyd 1960），直观地反映出不仅高比例的基因被转录，而且在绝大多数基因的 DNA 上密集存在的 RNA 聚合酶。这种特殊的转录状态在早期卵子发生中获得，并可能在成年雌性动物的卵巢中维持。成熟的精子，相反，被最大限度地压缩并且在转录水平上完全失活。常见的染色体组蛋白在精子中被精蛋白所替代。该蛋白在受精时进入卵的精子核中被交换，精子细胞核在 20 分钟内进行了非常快的去致密化过程。在两栖类中，没有发生哺乳动物中发生的 X 染色体失活和印记。在从受精到中-囊胚转化过程中（5 个小时）DNA 甲基化水平下降，在中-囊胚转化后，DNA 甲基化水平又随着发育过程逐渐上升（Meehan 2003）。综上所述，在正常发育过程中，基本的核重编程时间发生于配子发生和受精后的几个小时内。

两栖类核移植中发生的最明显的变化是染色质的弥散和体积增加。该现象在胚胎细胞来源的核中比分化的或成体细胞来源的核中发生的更快。在各方面，被移植的核适应了受精卵或卵细胞核的环境。核酸合成也跟随着核移植而变化。不分裂的细胞如成体脑细胞的细胞核在受精卵中很快发生 DNA 合成。从单个核移植的卵中获得的胚胎，合成了与正常受精发育而来的胚胎自身细胞核合成的水平相当的 rRNA 和 tRNA。基因转录模式从供体细胞模式转变成早期胚胎模式。例如，所有的基因转录在核移植胚胎的卵裂过程中被关闭，然后根据细胞类型在存活的核移植胚胎中被重新激活。即使核移植胚胎由小肠细胞核获得，肌肉形成相关基因也在核移植胚胎的肌肉细胞中表达（Gurdon et al. 1984）。

在核移植入卵的过程中，在没有任何 DNA 复制的情况下，被植入的细胞核发生了大规模的转录变化。例如，爪蟾肾细胞来源的核沉默了肾特异的基因，活化了卵特异的基因。一些新活化的基因是胚胎特异的，如小鼠胸腺细胞核在核移植后沉默了胸腺特异基因 *Thy-1* 而表达干细胞标记基因 *Oct-4*（Byrne et al. 2003）。总而言之，两栖类核移植入受精卵或卵子表现了广泛的基因转录的重编程，从而使得体细胞的核（以及在核移植入卵的例子中，它们的有丝分裂的后代）改变了其基因转录模式以适应于受体细胞的

转录模式。

### 在克隆中的重编程

在两栖类核移植实验中，长期以来认为由分化程度更高的供体细胞核来源的克隆胚胎发育异常的原因是不完全的 DNA 复制。在正常爪蟾发育过程中，雌雄原核染色体在受精后 20 分钟开始复制，并于 20 分钟内完成。相反，体外培养的细胞核分裂需要约 6 个小时来完成一次 DNA 复制。因此，移植后的体细胞核通常用远超过 40 分钟的时间继续 DNA 复制，直到第一次有丝分裂染色体凝聚并不奇怪。因此，染色体复制可能是不完全的，并且不完全复制的染色体被强行分开，即植入的核被强迫进入他们的第一次有丝分裂。破碎的染色体片段常在核移植胚胎中被发现（Di Berardino and Hoffner 1970），并且与体细胞相比，这种发生于受精卵的 DNA 复制速率和细胞分裂的不吻合性导致了染色体非整倍化，从而比较可能导致许多核移植胚胎的发育异常，尤其是高比例的核移植受精卵难以进行任何正常分裂。这在从非分裂的已分化细胞核来源的合子中可达到 75%。值得注意的是从部分分裂的第一次核移植胚胎中进行的系列核移植通常也可得到正常发育的蝌蚪。针对这一点的一个解释是体细胞核在卵的有丝分裂期提取物中可以很大程度地增加 DNA 复制起始位点的数量，从而使这些核比终末分化细胞（如红细胞）的核能更快地完成染色体复制（Lemaitre et al. 2005）。

另有两个解释可以帮助我们了解在中-囊胚转换（mid-blastula transition）中出现于合子转录起始的核移植发育异常。一个是早期合子基因活化在定量上的不规则性（Byrne et al. 2003），另一个是核移植胚胎中异常的生殖系供体特异性基因的持续表达。但是，还没有证据表明这些与正常基因表达的不同是导致所观察的发育异常的直接原因。

### 重编程的机制

由于两栖类受精卵和卵子数量较多且体积较大，为人们了解重编程分子机制提供了方便。最初的路线是获得可以体外模拟核移植入受精卵或卵子后所发生事件的细胞提取物。对提取物组分的去除可以帮助确定重编程所必需的组分。这个方法在确定 DNA 合成起始相关的卵的组分上尤为成功。核质的确定也值得注意（Laskey et al. 1978；Philpott et al. 1991），这是爪蟾卵中精子去致密化并促进组蛋白交换的一种高丰度物质。将体细胞核加入卵提取物中时，同样的过程也会发生（Dimitrov and Wolffe 1996；Tamada et al. 2006）。其他可能在核移植过程中起作用的卵提取物的组分包括染色体重塑复合物 ISWI（Kikyo et al. 2000）和可以可逆地解聚核仁的生殖细胞特异性蛋白 FRGY2（Gonda et al. 2003）。有人认为通过将提取物中的核通透和重封闭，染色质重塑复合物 BRG-1 也可能在卵和早期胚胎中起作用（Hansis et al. 2004）。由于无细胞提取物并不能起始核的转录，所以这些实验难以诠释。因此，在体外对核进行处理，然后将核植入活的卵子并来检测转录是所能做到的最有说服力的尝试（Byrne et al. 2003；Tamada et al. 2006）。

目前，可以认为成功的重编程需要三个必需步骤的：①在 DNA 或蛋白上的决定分化状态的表观遗传标记已去除；②新表达的那些基因所必需的转录因子；③相互作用基

因的染色质已去致密化使转录因子可以接近。

## 4.2　哺乳动物

连续的表观遗传重编程是正常发育的一个重要方面（Rideout et al. 2001）。在配子发生过程中，DNA 和组蛋白甲基化状态的变化连续发生于两个亲本基因组上。受精后，胚胎的基因组在卵裂过程中和植入子宫后又进一步被修饰。表 22-3 总结了克隆动物由于错误的重编程而区别于正常动物的一些表观遗传差异。在接下来的讨论中，我们关注正常受精个体和克隆个体在发育不同阶段的表观遗传差异。表 22-3 中描述的发育阶段按顺序为①配子发生，②卵裂，③植入后胚胎，④出生后发育。

**表 22-3　正常胚胎和克隆胚胎对比**

| 发育阶段 | 正常胚胎 | 克隆胚胎 |
|---|---|---|
| 配子发生 | 基因组足够活化“早期”基因，确立印记 | 无 |
| 卵裂期 | 全基因组范围的 DNA 去甲基化 | DNA 异常甲基化 |
| | 胚胎（“*Oct4* 样”）的基因活化 | “*Oct4* 样”的基因随机的/错误的活化 |
| 植入后 | 端粒长度调整 | 端粒长度正常 |
| | 基因组范围的 DNA 从头甲基化，X 染色体失活 | 某些克隆胚胎中异常 |
| | 正常印记和基因表达 | 异常印记，基因组范围的基因调节异常 |
| 出生后 | 正常的动物 | 大子代综合征，早熟死亡等 |

### 配子发生

正常胚胎发育中最重要的表观遗传重编程发生在配子发生时期。这个过程是指将精子和卵子基因组赋予“表观遗传感受态”，从而使今后的受精过程和早期发育关键基因的正常活化得以顺利进行（Latham 1999）。在克隆个体中，这个过程被截短了，并且许多影响克隆动物“正常化”的问题可能归因于植入卵中的体细胞核不适当的重编程。由于胎盘来自于胚胎中最先分化的细胞类型——滋养外胚层，一个可能的推测是在绝大多数克隆动物中，重编程和分化到这个早期细胞世系的过程受到了影响。确实，如下文总结的那样，在新生的克隆动物胎盘中，异常表达的基因比成体组织中明显要高。

### 卵裂

在卵裂过程中，基因组范围的去甲基化浪潮移去了存在于受精卵中的表观遗传标记，使得囊胚的 DNA 大部分没有甲基化。在植入子宫到原肠形成过程中，基因组范围的从头甲基化浪潮重新确立了所有的甲基化模式，这个模式随后贯穿动物体细胞的一生。在克隆胚胎中，重复序列的甲基化模式是异常的（Bourc'his et al. 2001；Dean et al. 2001；Kang et al. 2003；Mann et al. 2003）。为了研究基因表达，在克隆胚胎中检测到的细胞中沉默而在胚胎细胞中活化的“多能基因”如 *Oct-4* 的活性。突出的是，在大部分体 *Oct-4* 和“*Oct-4* 样”基因的重新活化是错误和随机的（Boiani et al. 2002；Bortvin et al. 2003）。因为缺乏 *Oct-4* 的胚胎会在发育中很早停滞，所以克隆个体中

"*Oct-4* 样"基因的不完全活化可能会导致绝大多数核移植胚胎无法活到植入后阶段。并且，一些研究已经检测到了克隆胚胎中异常的 DNA 甲基化。在核移植实验中需要在何种程度上模拟正常发育中染色质结构和 DNA 甲基化的表观遗传修饰才能成功，是一个未解决的问题，但现有的证据与错误的表观遗传重编程会导致克隆动物基因表达异常这一结论完全一致。

植入后发育

在植入后和原肠发生前，三个关键事件塑造了胚胎基因组的表观遗传状态：①基因组水平的从头甲基化重新确立了成体特异的并在体细胞整个生命过程中维持的全部甲基化模式（Dean et al. 2003）；②在雌性胚胎中通过随机失活两条 X 染色体中的一条而实现剂量补偿效应；③端粒调整为体细胞特异的长度。因为所有这些事件都在形成合子后才被起始，所以在克隆动物中并不预期发现表观遗传事件调节的紊乱。但是，在克隆牛胚胎而非新生牛胚胎中可以观察到基因组水平的低甲基化（Cezar et al. 2003）。在健康的克隆小鼠胚胎中 X 染色体失活是随机的且不受影响，但在异常发育的克隆小鼠胚胎中则并非如此（Eggan et al. 2000；Senda et al. 2004；Nolen et al. 2005）。然而这些紊乱是异常克隆发育的原因还是异常重编程过程在植入前发育的结果仍不清楚。相反，端粒长度的调整在克隆牛和小鼠中被准确完成（Lanza et al. 2000；Tian et al. 2000；Wakayama et al. 2000；Betts et al. 2001），因而应该不会影响克隆动物的存活。

出生后发育

在新生的克隆小鼠中对基因表达进行了最为广泛的研究。表达谱显示在新生克隆小鼠的胎盘中有 4%～5%的基因组和 30%～50%的印记基因是异常表达的（Humpherys et al. 2002；Kohda et al. 2005）。这说明了哺乳动物发育对于广泛的基因调节异常具有令人惊讶的忍耐性，而补偿机制确保了一些克隆可以存活到出生后。这些结果表明存活的克隆动物也可能有潜在的缺陷，虽然不会严重到威胁当前的存活，但仍可能在后期引发异常的表型。

## 5. 表观遗传记忆

目前已知的在脊椎动物发育过程中发生的基因组表观遗传修饰有两种。一是 DNA 中许多 CpG 二核苷酸存在的位点胞嘧啶的甲基化，一是组蛋白尾部的许多修饰。这些变化是在配子发生和早期发育中获得的，并且与基因的活化和失活密切相关。因此，可以推测这些表观遗传修饰可以通过核移植被逆转。如果不能被逆转，那么它们有助于了解一些核移植胚胎发育失败的原因。

在两栖类中，可以通过将哺乳动物体细胞核注射入爪蟾卵的实验来研究 DNA 去甲基化（Simonsson and Gurdon 2004）。小鼠 *Oct4* 基因的启动子在成体胸腺细胞中被甲基化而不表达。但是，当胸腺细胞核被注射入卵中，该基因调节区域的启动子而非增强子

区被去甲基化。这个结果显示了去甲基化过程的选择性。当完整的核被注入卵细胞时，DNA 去甲基化似乎先于 *Oct4* 转录的诱导。DNA 去甲基化酶活性很可能是卵所特有的（见上文），并且发育中被抑制的基因的启动子 DNA 的去甲基化也可能是体细胞核移植重编程中重要且必须的步骤。组蛋白修饰的变化还未在两栖类核移植实验中检测过。

两栖类核移植实验的另一个设计已经说明体细胞的表观遗传状态并不总能被逆转。纵观核移植失败的例子，甚至从来自 *R. pipiens* 早期的尾端内胚层供体开始，Briggs 和 Kings（1957）就提出异常发育胚胎的形态是否反映它们来源的问题。他们描述了供体核来源于内胚层的克隆胚胎中内胚层组织的存活更加优先，并称之为“内胚层症状”。但是，需要指出的是，内胚层的分化晚于其他胚层，这可能是它能更好存活的原因（Gurdon 1963）。确实，Simnett（1964）报道了神经细胞来源的核移植胚胎也表现了相同的内胚层优先存活。运用细胞类型特异的基因标记，相同的问题在最近又被提出。在这些实验中（Ng and Gurdon 2005），来自神经外胚层或内胚层的分别表达细胞类型特异标记 *Sox2* 或 *endodermin* 的细胞核，被植入去核的卵中；得到的核移植胚胎被分成神经外胚层或内胚层部分，并且用相同的 *Sox2* 或 *endodermin* 标记进行检测（图 22-8 和表 22-4）。结果发现两个基因标记都在不正常的细胞类型中被优先表达。例如，有超过一半的神经外胚层来源的克隆胚胎在其内胚层细胞中过表达神经标记 *Sox2*。在一些克隆胚胎中，*Sox2* 基因表达水平看起来并不比供体细胞中的表达水平低。与那些正常胚胎类似，被移植的核在转录水平整体失活直到囊胚时期。因此，在细胞分化过程中建立的基因转录活化状态可以在克隆胚胎中被维持是显而易见的，无需通过 12 个有丝分裂细胞周期（从受精卵到囊胚）来对其进行诱导。该表观遗传记忆的明显的例子在部分核移植胚胎中可以看到，但在基因表达被成功重编程的其他克隆胚胎中则完全没有。

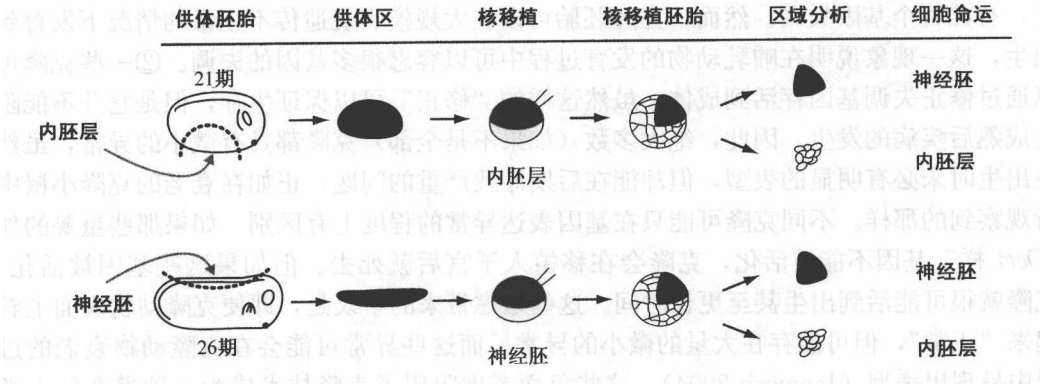



图 22-8　检测细胞类型特异的基因表达的表观遗传记忆的实验设计

由来源于 21 期的内胚层或 26 期的神经细胞得到的核移植胚胎通常形成部分囊胚，其正常分裂的部分分化为将来的神经胚区域或内胚层区域，对他们基因表达分析的结果见表 22-4（经许可重印自 Ng and Gurdon 2005）。

表 22-4 细胞类型特异的基因表达的表观遗传印记

| | | 基因表达(%) | |
|---|---|---|---|
| | | *Edd* | *Sox2* |
| 内胚层核($edd^+$)→核移植胚胎 | 神经外胚层 | 45 | 5 |
| | 内胚层 | 12 | 0 |
| 神经胚层核($Sox2^+$)→核移植胚胎 | 神经外胚层 | 6 | 22 |
| | 内胚层 | 0 | 81 |

注:数值代表了 *Edd* 或 *Sox2* 基因表达高于正常(或背景)水平两倍或以上的核移植胚胎的比例(通过单独 RT-PCR 确定)(数据来源于 Ng and Gurdon 2005)。

## 6. 核移植的医学应用

区分“繁殖性克隆”和“核移植治疗”(也称治疗性克隆或 SCNT)是十分重要的。在繁殖性克隆中,通过移植体细胞的细胞核到去核的卵细胞中,目的是能够创造一个新的克隆个体。而核移植治疗的目的是通过患者提供的细胞核构建一个胚胎干细胞系(也叫 ntES 细胞),并按照患者的需要来对 ES 进行“剪裁”(Hochedlinger and Jaenisch 2003)。ntES 作为原材料,为用于治疗某些重要疾病的功能细胞提供材料。图22-9中列出了正常受精卵的发育,繁殖性克隆和治疗性克隆的比较。

### 6.1 繁殖性克隆

在前面的概述中,从已经克隆的八种哺乳动物中可以看出,通过克隆技术得到正常个体的过程还面临着巨大的障碍。其中大众关心的一个重要的疑问是通过核移植得到正常的个体是否可行。从已知的证据可以看出,正常的克隆如果不是完全不能得到,至少也是非常困难的,原因如下:①如上所述,所有克隆胚胎在初生的时候与正常胚胎相比,有几百个基因失调。然而,克隆胚胎可以在大规模表观遗传不正常的情况下发育到出生,这一现象说明在哺乳动物的发育过程中可以容忍很多基因的失调。②一些克隆可以通过修正失调基因存活到成体。虽然这样的“修正”可以保证生存,但是它并不能阻止成熟后疾病的发生。因此,绝大多数(如果不是全部)克隆都具有微小的异常,虽然在出生时未必有明显的表型,但却能在后期导致严重的问题,正如在衰老的克隆小鼠中所观察到的那样。不同克隆可能只在基因表达异常的程度上有区别。如果那些重要的如“*Oct4* 样”基因不能被活化,克隆会在移植入子宫后就死去。但如果这些基因被活化,克隆就很可能活到出生甚至更长时间。这些考虑带来的争议是,即便克隆动物表面上看起来“正常”,但可能存在大量的微小的异常,而这些异常可能会在克隆动物衰老的过程中呈现出表型(Jaenisch 2004)。这些负面考虑阻碍了克隆技术成为一种潜在的人类繁殖技术。

### 6.2 核移植的治疗性应用

由于免疫不相容性,异体器官移植通常会发生免疫排斥。为了治疗这种“宿主针对

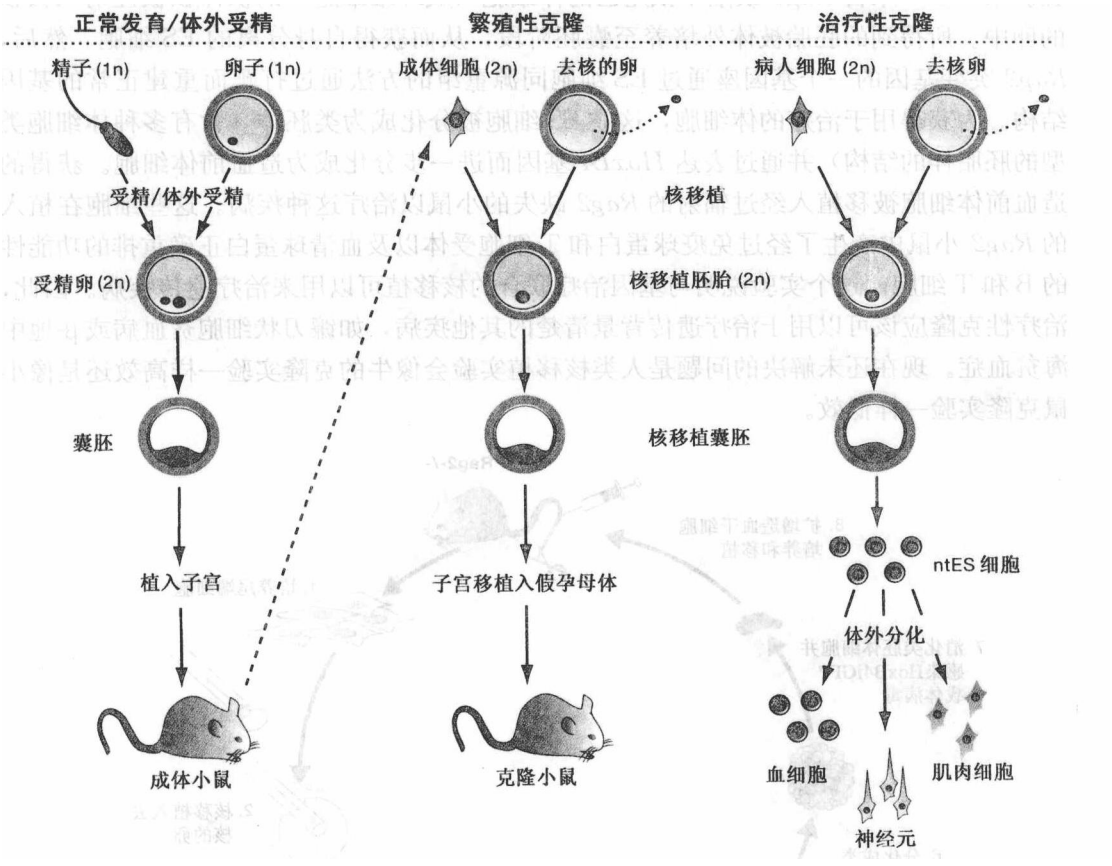



图 22-9　“繁殖性克隆”和“治疗性克隆”与正常胚胎发育的比较

（左图）在正常发育中，一个单倍体的精子与一个单倍体的卵子受精形成一个 2 倍体（2n）的合子，合子进行卵裂形成囊胚。囊胚植入子宫并最终发育成新生动物。（中图）在“繁殖性克隆”中，一个 2 倍体成体供体细胞核被移入去核的受体卵中，在人工激活后，卵裂成为一个克隆囊胚。在植入假孕母体后，一些克隆的囊胚可发育成新生动物。（右图）相反，通过核移植获得 ntES 细胞需要将克隆胚胎体外种植培养。该 ES 细胞系可被体外分化成为任何细胞类型从而用于研究或治疗（经许可重印自 Hochedlinger and Jaenisch 2003 [Massachusets Medical Society]）。

移植物”的疾病，移植受体按惯例接受免疫抑制药物，而这种治疗有严重的副作用。由核移植来源的胚胎干细胞与患者细胞遗传上完全相同，因而消除了免疫排斥的风险和对免疫抑制的需求。更重要的是，为治疗遭受严重紊乱疾病如帕金森病、心脏病或糖尿病的患者而生成功能性细胞如神经原、肌肉细胞和胰岛细胞的方法也被发展出来。并且，胚胎干细胞为组织替代提供了一个可更新的源泉，从而在有必要的情况下使重复治疗成为可能。

的确，治疗性克隆的可行性已在疾病动物模型中被验证。在这里，“患者”是一个 *Rag2* 基因缺失的小鼠品系（Rideout et al. 2002）。*Rag2* 突变的小鼠由于催化淋巴细胞免疫受体重排的基因突变而患有严重的综合免疫缺陷（SCID）。这些小鼠缺乏成熟的 B 和 T 细胞，类似于人的一种疾病（“安全室婴儿”）。图 22-10 总结了这个实验需要的步

骤。第一步，来自 *Rag2* 缺陷小鼠尾巴的体细胞（成纤维细胞）的供体核被注射入去核的卵中。所得到的胚胎被体外培养至囊胚阶段，从而获得自身分离的 ES 细胞。然后，*Rag2* 突变基因的一个基因座通过 ES 细胞同源重组的方法通过打靶而重建正常的基因结构。为获得用于治疗的体细胞，这些 ES 细胞被分化成为类胚体（含有多种体细胞类型的胚胎样的结构）并通过表达 *HoxB4* 基因而进一步分化成为造血前体细胞。获得的造血前体细胞被移植入经过辐射的 *Rag2* 缺失的小鼠以治疗这种疾病。这些细胞在植入的 *Rag2* 小鼠中产生了经过免疫球蛋白和 T 细胞受体以及血清球蛋白正确重排的功能性的 B 和 T 细胞。这个实验说明与基因治疗联合的核移植可以用来治疗遗传疾病。因此，治疗性克隆应该可以用于治疗遗传背景清楚的其他疾病，如镰刀状细胞贫血病或 β-地中海贫血症。现在还未解决的问题是人类核移植实验会像牛的克隆实验一样高效还是像小鼠克隆实验一样低效。

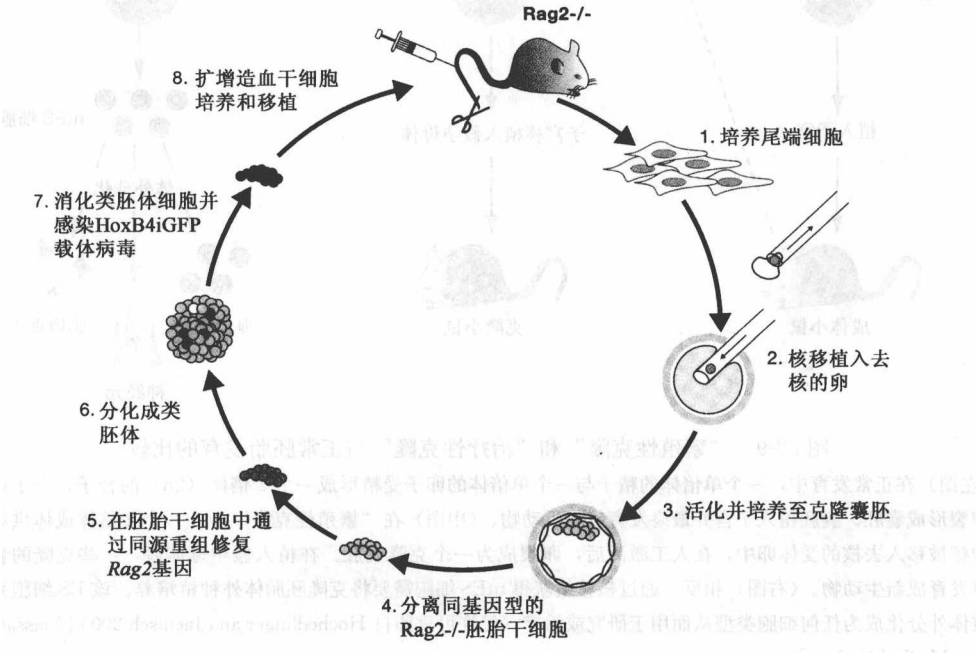



图 22-10 结合基因和细胞治疗的治疗性克隆流程图

来自重组活化基因 2（*Rag2*）突变的纯合小鼠尾部进行细胞培养。当成纤维细胞长出后，它们被用作供体核通过一个 Piezoelectirc 推动的显微操作杆直接注入去核的 MII 卵中。由克隆囊胚获得的胚胎干细胞（ES）通过同源重组被遗传修复。修复后，ntES 细胞被体外分化成类胚体（EB），再经 HoxB4iGFP 逆转录病毒感染，扩增后植入经辐射的 *Rag2* 缺陷的小鼠尾静脉中（经许可重印自 Hochedlinger and Jaenisch 2003 [© Massachusetts Medical Society]）。

运用核移植来“定制”ES 细胞仍有争议，因为不清楚这样产生的 ES 细胞是否破坏了潜在的人类生命。作为解决这个伦理难题的一个可行的方案，改版核移植（ANT）被建议作为核移植过程的一个修正（Hurlbut 2005）。ANT 指将一个在胎盘发育中必要基因进行失活，使囊胚不能着床从而阻止了胎儿的形成，因此不会破坏潜在的人类生命，但仍能“定制”ES 细胞。将 *Cdx2* 作为目标基因，在小鼠中进行的实验证明了

ANT 方法的可行性（Meissner and Jaenisch 2006）。但 ANT 修饰是否能够满足那些反对从克隆人胚胎中获得 ES 细胞的人，值得进一步关注。

## 6.3　繁殖性克隆与治疗性克隆的区别

为何重编程问题困扰着繁殖性克隆而非治疗性克隆呢？对于这个看起来自相矛盾的问题，最重要的原因是：与繁殖性克隆相反，治疗性克隆不需要形成胎儿而只是在体外定向分化为功能性的细胞。因为不需要发育成胎儿，这些分化形成的细胞的功能不会受到会导致发育失败的异常的印记基因的影响（Jaenisch 2004）。因为来源于正常受精胚胎的 ES 细胞可以形成所有正常的胚胎组织，通过核移植得到的 ES 细胞也应该具有相似的潜能。的确，所有可行的证据与克隆胚胎来源的 ES 细胞在功能和分子水平上与受精胚胎来源的 ES 细胞相似的结论一致（Brambrink et al. 2006）。因此，如果通过 IVF 胚胎获得的人 ES 细胞可用来治疗疾病，那么通过核移植获得患者特有的“定做的”ES 细胞也可以用来治疗疾病。

王译萱　杨　鹏　译
马　奔　校
朱　冰　审校

## 参考文献

Betts D., Bordignon V., Hill J., Winger Q., Westhusin M., Smith L., and King W. 2001. Reprogramming of telomerase activity and rebuilding of telomere length in cloned cattle. *Proc. Natl. Acad. Sci.* **98:** 1077–1082.

Blelloch R.H., Hochedlinger K., Yamada Y., Brennan C., Kim M., Mintz B., Chin L., and Jaenisch R. 2004. Nuclear cloning of embryonal carcinoma cells. *Proc. Natl. Acad. Sci.* **101:** 13985–13990.

Boiani M., Eckardt S., Scholer H.R., and McLaughlin K.J. 2002. Oct4 distribution and level in mouse clones: Consequences for pluripotency. *Genes Dev.* **16:** 1209–1219.

Bolton V.N., Oades P.J., and Johnson M.H. 1984. The relationship between cleavage, DNA replication, and gene expression in the mouse 2-cell embryo. *J. Embryol. Exp. Morphol.* **79:** 139–163.

Bortvin A., Eggan K., Skaletsky H., Akutsu H., Berry D.L., Yanagimachi R., Page D.C., and Jaenisch R. 2003. Incomplete reactivation of *Oct4*-related genes in mouse embryos cloned from somatic nuclei. *Development* **130:** 1673–1680.

Bourc'his D., Le Bourhis D., Patin D., Niveleau A., Comizzoli P., Renard J.P., and Viegas-Pequignot E. 2001. Delayed and incomplete reprogramming of chromosome methylation patterns in bovine cloned embryos. *Curr. Biol.* **11:** 1542–1546.

Brambrink T., Hochedlinger K., Bell G., and Jaenisch R. 2006. ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable. *Proc. Natl. Acad. Sci.* **103:** 933–938.

Briggs R. and King T.J. 1952. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proc. Natl. Acad. Sci.* **38:** 455–463.

———. 1957. Changes in the nuclei of differentiating endoderm cells as revealed by nuclear transplantation. *J. Morphol.* **100:** 269–312.

Bromhall J.D. 1975. Nuclear transplantation in the rabbit egg. *Nature* **258:** 719–722.

Byrne J.A., Simonsson S., Western P.S., and Gurdon J.B. 2003. Nuclei of adult mammalian somatic cells are directly reprogrammed to *oct-4* stem cell gene expression by amphibian oocytes. *Curr. Biol.* **13:** 1206–1213.

Calarco P.G. and McLaren A. 1976. Ultrastructural observations of preimplantation stages of the sheep. *J. Embryol. Exp. Morphol.* **36:** 609–622.

Callan H.G. and Lloyd L. 1960. Lampbrush chromosomes of crested newts *Triturus cristatus* (Laurenti). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **243:** 135–219.

Camous S., Kopecny V., and Flechon J.E. 1986. Autoradiographic detection of the earliest stage of [3H]-uridine incorporation into the cow embryo. *Biol. Cell* **58:** 195–200.

Campbell K.H., Loi P., Otaegui P.J., and Wilmut I. 1996a. Cell cycle coordination in embryo cloning by nuclear transfer. *Rev. Reprod.* **1:** 40–46.

Campbell K.H., McWhir J., Ritchie W.A., and Wilmut I. 1996b. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380:** 64–66.

Campbell K.H., Alberio R., Choi I., Fisher P., Kelly R.D., Lee J.H., and Maalouf W. 2005. Cloning: Eight years after Dolly. *Reprod. Domest. Anim.* **40:** 256–268.

Cezar G.G., Bartolomei M.S., Forsberg E.J., First N.L., Bishop M.D., and Eilertsen K.J. 2003. Genome-wide epigenetic alterations in cloned bovine fetuses. *Biol. Reprod.* **68:** 1009–1014.

Cibelli J.B., Stice S.L., Golueke P.J., Kane J.J., Jerry J., Blackwell C., Ponce de Leon F. A., and Robl J.M. 1998. Cloned transgenic calves produced from nonquiescent fetal fibroblasts. *Science* **280:** 1256–1258.

Dean W., Santos F., and Reik W. 2003. Epigenetic reprogramming in early mammalian development and following somatic nuclear transfer. *Semin. Cell Dev. Biol.* **14:** 93–100.

Dean W., Santos F., Stojkovic M., Zakhartchenko V., Walter J., Wolf E., and Reik W. 2001. Conservation of methylation reprogramming in mammalian development: Aberrant reprogramming in cloned embryos. *Proc. Natl. Acad. Sci.* **98:** 13734–13738.

Di Berardino M.A. and Hoffner N. 1970. Origin of chromosomal abnormalities in nuclear transplants—A reevaluation of nuclear differentiation and nuclear equivalence in amphibians. *Dev. Biol.*

**23:** 185–209.

Dimitrov S. and Wolffe A.P. 1996. Remodeling somatic nuclei in *Xenopus laevis* egg extracts: Molecular mechanisms for the selective release of histones H1 and H1° from chromatin and the acquisition of transcriptional competence. *EMBO J.* **15:** 5897–5906.

Eggan K., Akutsu H., Hochedlinger K., Rideout W., Yanagimachi R., and Jaenisch R. 2000. X-Chromosome inactivation in cloned mouse embryos. *Science* **290:** 1578–1581.

Eggan K., Akutsu H., Loring J., Jackson-Grusby L., Klemm M., Rideout W.M., III, Yanagimachi R., and Jaenisch R. 2001. Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. *Proc. Natl. Acad. Sci.* **98:** 6209–6214.

Eggan K., Baldwin K., Tackett M., Osborne J., Gogos J., Chess A., Axel R., and Jaenisch R. 2004. Mice cloned from olfactory sensory neurons. *Nature* **428:** 44–49.

Fischberg M., Gurdon J.B., and Elsdale T.R. 1958. Nuclear transplantation in *Xenopus laevis*. *Nature* **181:** 424.

Gonda K., Fowler J., Katoku-Kikyo N., Haroldson J., Wudel J., and Kikyo N. 2003. Reversible disassembly of somatic nucleoli by the germ cell proteins FRGY2a and FRGY2b. *Nat. Cell Biol.* **5:** 205–210.

Gurdon J.B. 1960. The developmental capacity of nuclei taken from differentiating endoderm cells of *Xenopus laevis*. *J. Embryol. Exp. Morphol.* **8:** 505–526.

———. 1962. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J. Embryol Exp. Morphol.* **10:** 622–640.

———. 1963. Nuclear transplantation in Amphibia and the importance of stable nuclear changes in cellular differentiation. *Q. Rev. Biol.* **38:** 54–78.

———. 1964. The transplantation of living cell nuclei. *Adv. Morphog.* **4:** 1–43.

Gurdon J.B. and Uehlinger V. 1966. "Fertile" intestine nuclei. *Nature* **210:** 1240–1241.

Gurdon J.B., Elsdale T.R., and Fischberg M. 1958. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* **182:** 64–65.

Gurdon J.B., Brennan S., Fairman S., and Mohun T.J. 1984. Transcription of muscle-specific actin genes in early *Xenopus* development: Nuclear transplantation and cell dissociation. *Cell* **38:** 691–700.

Hansis C., Barreto G., Maltry N., and Niehrs C. 2004. Nuclear reprogramming of human somatic cells by *Xenopus* egg extract requires BRG1. *Curr. Biol.* **14:** 1475–1480.

Hill J.R., Burghardt R.C., Jones K., Long C.R., Looney C.R., Shin T., Spencer T.E., Thompson J.A., Winger Q.A., and Westhusin M.E. 2000. Evidence for placental abnormality as the major cause of mortality in first-trimester somatic cell cloned bovine fetuses. *Biol. Reprod.* **63:** 1787–1794.

Hochedlinger K. and Jaenisch R. 2002. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* **415:** 1035–1038.

———. 2003. Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. *N. Engl. J. Med.* **349:** 275–286.

Hochedlinger K., Blelloch R., Brennan C., Yamada Y., Kim M., Chin L., and Jaenisch R. 2004. Reprogramming of a melanoma genome by nuclear transplantation. *Genes Dev.* **18:** 1875–1885.

Humpherys D., Eggan K., Akutsu H., Friedman A., Hochedlinger K., Yanagimachi R., Lander E., Golub T.R., and Jaenisch R. 2002. Abnormal gene expression in cloned mice derived from embryonic stem cell and cumulus cell nuclei. *Proc. Natl. Acad. Sci.* **99:** 12889–12894.

Humpherys D., Eggan K., Akutsu H., Hochedlinger K., Rideout W., Biniszkiewicz D., Yanagimachi R., and Jaenisch R. 2001. Epigenetic instability in ES cells and cloned mice. *Science* **293:** 95–97.

Hurlbut W.B. 2005. Altered nuclear transfer as a morally acceptable means for the procurement of human embryonic stem cells. *Perspect. Biol. Med.* **48:** 211–228.

Inoue K., Wakao H., Ogonuki N., Miki H., Seino K., Nambu-Wakao R., Noda S., Miyoshi H., Koseki H., Taniguchi M., and Ogura A. 2005. Generation of cloned mice by direct nuclear transfer from natural killer T cells. *Curr. Biol.* **15:** 1114–1118.

Jaenisch R. 2004. Human cloning—The science and ethics of nuclear transplantation. *N. Engl. J. Med.* **351:** 2787–2791.

Kang Y.K., Lee K.K., and Han Y.M. 2003. Reprogramming DNA methylation in the preimplantation stage: Peeping with Dolly's eyes. *Curr. Opin. Cell Biol.* **15:** 290–295.

Kikyo N., Wade P.A., Guschin D., Ge H., and Wolffe A.P. 2000. Active remodeling of somatic nuclei in egg cytoplasm by the nucleosomal ATPase ISWI. *Science* **289:** 2360–2362.

Kohda T., Inoue K., Ogonuki N., Miki H., Naruse M., Kaneko-Ishino T., Ogura A., and Ishino F. 2005. Variation in gene expression and aberrantly regulated chromosome regions in cloned mice. *Biol. Reprod.* **73:** 1302–1311.

Lanza R., Cibelli J., Blackwell C., Cristofalo V., Francis M., Baerlocher G., Mak J., Schertzer M., Chavez E., Sawyer N., et al. 2000. Extension of cell life-span and telomere length in animals cloned from senescent somatic cells. *Science* **288:** 665–669.

Laskey R.A. and Gurdon J.B. 1970. Genetic content of adult somatic cells tested by nuclear transplantation from cultured cells. *Nature* **228:** 1332–1334.

Laskey R.A., Honda B.M., Mills A.D., and Finch J.T. 1978. Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature* **275:** 416–420.

Latham K.E. 1999. Mechanisms and control of embryonic genome activation in mammalian embryos. *Int. Rev. Cytol.* **193:** 71–124.

Lemaitre J.M., Danis E., Pasero P., Vassetzky Y., and Mechali M. 2005. Mitotic remodeling of the replicon and chromosome structure. *Cell* **123:** 787–801.

Li J., Ishii T., Feinstein P., and Mombaerts P. 2004. Odorant receptor gene choice is reset by nuclear transfer from mouse olfactory sensory neurons. *Nature* **428:** 393–399.

Li L., Connelly M.C., Wetmore C., Curran T., and Morgan J.I. 2003. Mouse embryos cloned from brain tumors. *Cancer Res.* **63:** 2733–2736.

Mann M.R., Chung Y.G., Nolen L.D., Verona R.I., Latham K.E., and Bartolomei M.S. 2003. Disruption of imprinted gene methylation and expression in cloned preimplantation stage mouse embryos. *Biol. Reprod.* **69:** 902–914.

McGrath J. and Solter D. 1984a. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37:** 179–187.

McGrath J. and Solter D. 1984b. Inability of mouse blastomere nuclei transferred to enucleated zygotes to support development in vitro. *Science* **226:** 1317–1319.

McKinnell R.G. 1962. Development of *Rana pipiens* eggs transplanted with Lucke tumor cells. *Am. Zool.* **2:** 430–431.

Meehan R.R. 2003. DNA methylation in animal development. *Semin. Cell Dev. Biol.* **14:** 53–65.

Meissner A. 2006. "Conditional RNA interference, altered nuclear transfer and genome-wide DNA methylation analysis." Ph.D. thesis. University Saarland, Saarbrücken, Germany.

Meissner A. and Jaenisch R. 2006. Generation of nuclear transfer-derived pluripotent ES cells from cloned *Cdx2*-deficient blastocysts. *Nature* **439:** 212–215.

Ng R.K. and Gurdon J.B. 2005. Epigenetic memory of active gene transcription is inherited through somatic cell nuclear transfer. *Proc. Natl. Acad. Sci.* **102:** 1957–1962.

Nolen L.D., Gao S., Han Z., Mann M.R., Gie Chung Y., Otte A.P., Bartolomei M.S., and Latham K.E. 2005. X chromosome reactivation and regulation in cloned embryos. *Dev. Biol.* **279:** 525–540.

Ogonuki N., Inoue K., Yamamoto Y., Noguchi Y., Tanemura K., Suzuki O., Nakayama H., Doi K., Ohtomo Y., Satoh M., et al. 2002. Early death of mice cloned from somatic cells. *Nat. Genet.* **30:** 253–254.

Ogura A., Inoue K., Ogonuki N., Noguchi A., Takano K., Nagano R., Suzuki O., Lee J., Ishino F., and Matsuda J. 2000. Production of male cloned mice from fresh, cultured, and cryopreserved immature Sertoli cells. *Biol. Reprod.* **62:** 1579–1584.

Philpott A., Leno G.H., and Laskey R.A. 1991. Sperm decondensation in *Xenopus* egg cytoplasm is mediated by nucleoplasmin. *Cell* **65:**

569–578.

Prather R.S., Sims M.M., and First N.L. 1989. Nuclear transplantation in early pig embryos. *Biol. Reprod.* **41:** 414–418.

Rideout W.M., III, Eggan K., and Jaenisch R. 2001. Nuclear cloning and epigenetic reprogramming of the genome. *Science* **293:** 1093–1098.

Rideout W.M., *III*, Hochedlinger K., Kyba M., Daley G.Q., and Jaenisch R. 2002. Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. *Cell* **109:** 17–27.

Rideout W.M., III, Wakayama T., Wutz A., Eggan K., Jackson-Grusby L., Dausman J., Yanagimachi R., and Jaenisch R. 2000. Generation of mice from wild-type and targeted ES cells by nuclear cloning. *Nat. Genet.* **24:** 109–110.

Robi J.M., Prather R., Barnes F., Eyestone W., Northey D., Gilligan B., and First N.L. 1987. Nuclear transplantation in bovine embryos. *J. Anim. Sci.* **64:** 642–647.

Senda S., Wakayama T., Yamazaki Y., Ohgane J., Hattori N., Tanaka S., Yanagimachi R., and Shiota K. 2004. Skewed X-inactivation in cloned mice. *Biochem. Biophys. Res. Commun.* **321:** 38–44.

Simnett J.D. 1964. The development of embryos derived from the transplantation of neural ectoderm cell nuclei in *Xenopus laevis*. *Dev. Biol.* **10:** 467–486.

Simonsson S. and Gurdon J. 2004. DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nat. Cell Biol.* **6:** 984–990.

Surani M.A., Barton S.C., and Norris M.L. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308:** 548–550.

Tamada H., Van Thuan N., Reed P., Nelson D., Katoku-Kikyo N., Wudel J., Wakayama T., and Kikyo N. 2006. Chromatin decondensation and nuclear reprogramming by nucleoplasmin. *Mol. Cell. Biol.* **26:** 1259–1271.

Tamashiro K.L., Wakayama T., Akutsu H., Yamazaki Y., Lachey J.L., Wortman M.D., Seeley R.J., D'Alessio D.A., Woods S.C., Yanagimachi R., and Sakai R.R. 2002. Cloned mice have an obese phenotype not transmitted to their offspring. *Nat. Med.* **8:** 262–267.

Tanaka S., Oda M., Toyoshima Y., Wakayama T., Tanaka M., Yoshida N., Hattori N., Ohgane J., Yanagimachi R., and Shiota K. 2001. Placentomegaly in cloned mouse concepti caused by expansion of the spongiotrophoblast layer. *Biol. Reprod.* **65:** 1813–1821.

Tian X.C., Xu J., and Yang X. 2000. Normal telomere lengths found in cloned cattle. *Nat. Genet.* **26:** 272–273.

Wakayama S., Ohta H., Kishigami S., Van Thuan N., Hikichi T., Mizutani E., Miyake M., and Wakayama T. 2005. Establishment of male and female nuclear transfer embryonic stem cell lines from different mouse strains and tissues. *Biol. Reprod.* **72:** 932–936.

Wakayama T. and Yanagimachi R. 1999. Cloning of male mice from adult tail-tip cells. *Nat. Genet.* **22:** 127–128.

———. 2001. Mouse cloning with nucleus donor cells of different age and type. *Mol. Reprod. Dev.* **58:** 376–383.

Wakayama T., Perry A.C., Zuccotti M., Johnson K.R., and Yanagimachi R. 1998. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* **394:** 369–374.

Wakayama T., Rodriguez I., Perry A.C., Yanagimachi R., and Mombaerts P. 1999. Mice cloned from embryonic stem cells. *Proc. Natl. Acad. Sci.* **96:** 14984–14989.

Wakayama T., Shinkai Y., Tamashiro K.L., Niida H., Blanchard D.C., Blanchard R.J., Ogura A., Tanemura K., Tachibana M., Perry A.C., et al. 2000. Cloning of mice to six generations. *Nature* **407:** 318–319.

Wilmut I., Schnieke A.E., McWhir J., Kind A.J., and Campbell K.H. 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385:** 810–813.

Young L.E., Sinclair K.D., and Wilmut I. 1998. Large offspring syndrome in cattle and sheep. *Rev. Reprod.* **3:** 155–163.

# 第 23 章

# 表观遗传学和人类疾病

Huda Y. Zoghbi[1] and Arthur L. Beaudet[2]

*[1]Howard Hughes Medical Institute, and [2]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030*

## 本章目录

1. 前言
2. 人类病例研究揭示表观遗传的生物学作用
3. 人类疾病
   - 3.1 基因组印记紊乱
   - 3.2 反式作用于染色质结构的紊乱
   - 3.3 顺式作用于染色质结构的紊乱
   - 3.4 表观遗传与环境的相互作用
4. 展望未来

致谢

参考文献

## 概要

在过去的二十年里，我们见证了在人类疾病遗传学基础认识上取得的前所未有的成功。对基因型与表型之间关系的研究对医生和科研人员提出了挑战，因为有一些观察到的现象无法用已有的知识很好地解释。例如，携带相同致病突变的同卵双胞胎可以有相当不同的临床表现。一个相同的突变由于其来自父本或母本的不同，可以在同一个家族子代中产生疾病上的巨大的差异。这些特殊病例的研究揭示了表观遗传组（在不改变DNA序列的情况下改变遗传信息）在人类健康和疾病中所起的作用。这些研究揭示了在哺乳动物某些基因组区域，父本与母本等位基因的功能并非相同。有的患者的两条同源染色体（或其片段）均来自于从同一亲本——单亲源二倍体（UPD）——他们失去了母源性特异表达基因（对父源性二倍体来说）的表达能力，却高表达父源性特异表达的基因。UPD以及DNA修饰（可改变DNA甲基化的表观遗传突变改变）很快被人们认识到是许多不同发育和神经疾病的分子基础。有趣的是对于许多这类疾病，表观遗传和遗传突变都可以导致相同的表型。这通常是因为遗传突变发生于常为表观遗传缺陷所影响的基因座。

另一类疾病中遗传突变导致一些参与DNA甲基化和染色质重塑的蛋白失去功能，其表型由一个或多个基因座的表观遗传状态改变引起。基因组和表观基因组间的关系拓展了致病性分子事件的类型。它们可以是继承性的或获得性的，遗传的或表观遗传的，最有趣的是，其中有些会受环境因素的影响。研究发现环境因素如食物，经历会改变表观基因组（尤其是DNA甲基化）。这给我们提供了一个新的机制来解释具遗传倾向性，又受环境影响的疾病，包括神经管缺陷和心理疾病。对影响表观基因组的环境因子的鉴定，有助发展介入疗法以降低疾病发生率或减轻其症状，如发育异常、癌症以及神经性心理疾病。

## 1. 前　言

两个遗传上相同的雄性同卵双生双胞胎，成长于相同的环境，却表现出了非常不同的神经生物学表型。双胞胎在 X 连锁的肾上腺脑白质营养不良（*ALD*）基因上有相同的突变，然而其中一个人失明，不能平衡，大脑中缺少髓磷脂，典型的致死神经发育疾病特点。而另一个则非常健康。对于这种非正常现象，研究人员的结论是“可能有一些非遗传性因素对不同的肾上腺脑白质营养不良表型起到了非常重要的作用（Korenke et al. 1996）。考虑到当时医学遗传学的焦点是 DNA 序列，在 1996 年就能得出这一结论是非常准确的。如果 DNA 序列不能解释表型差异，那么环境因素应该可以。与肾上腺脑白质营养不良的双胞胎类似，有些同卵双胞胎尽管在相似的环境条件下长大，却有不同的精神分裂症表型（Petronis 2004）。庆幸的是，过去的十几年里科学研究终于聚焦于表观遗传变化，即不改变 DNA 序列的条件下的遗传信息修饰。这对拥有相同 DNA 序列却呈现不同表型的同卵双胞胎提供了一种可能的解释（Dennis 2003；Fraga et al. 2005）。

表观遗传修饰控制基因在细胞中的表达模式。这些修饰稳定并可继承，所以一个肝脏细胞在分裂后产生的是更多的子代肝脏细胞。对于那些不分裂的细胞（如神经细胞），染色体特定区域上的染色质修饰为维持表观遗传信息提供了一种机制，还有可能介导了神经细胞对特定刺激反应的可重复性。表观遗传基因型（epigenotype，基因座的表观遗传状态）由 DNA 甲基化状态、染色质修饰以及尚有待明晰的非编码 RNA 的一系列活性所建立。

在哺乳动物中，迄今研究最透彻的表观遗传信号——DNA 甲基化主要发生在对称的 CpG 双核苷酸上胞嘧啶的第五位碳上。细胞分裂后，DNA 甲基化状态的维持依赖于 DNA 甲基转移酶 I，它甲基化子代细胞中半甲基化的 CpG 双核苷酸。染色质修饰包括组蛋白 N 端尾巴的乙酰化、甲基化、磷酸化、泛素化等共价转录后修饰。甲基化修饰可以分为单甲基化，双甲基化，三甲基化。这些修饰组成了可能的“组蛋白密码”，代表特定的染色质结构，并继而影响邻近基因的表达。由于染色质由高度折叠的，缠绕在组蛋白上的 DNA 组成，DNA 折叠进染色质的模式显然与基因活性变化紧密相关。尽管组蛋白密码和染色质结构可以稳定的从母代细胞传递到子代细胞，但是其复制机制尚未被充分理解。在发育中和出生后，由于环境以及经历等因素的影响，表观遗传型也展示出一定的可塑性。因此，就不难理解表观遗传型不仅能影响人类的发育疾病，还可以影响出生后甚至成年人的疾病。最新发现的一类对表观遗传信号有影响的分子是非编码 RNA。长期以来，非蛋白质编码 RNA 包括 tRNA，rRNA，spliceosomal RNA。最近，由于获得了许多生物的基因组序列，再通过物种间（从大肠杆菌到人类）分子遗传学研究。非编码 RNA 的种类得到了很大的扩充，发现了核仁小 RNA（snoRNA）、微 RNA（microRNA，miRNA），小干扰 RNA（small interfer RNA，siRNA）和双链小 RNA（small double-stranded RNA）。某些这类小 RNA 分子可调控染色质修饰，基因组印记，DNA 甲基化和转录沉默（详见第八章）。

第一个证明表观遗传在人类疾病当中起作用的明确证据来源于人们对基因组印记的理解以及发现若干基因受到基因组印记的调节（Reik 1989）。基因组印记是一种表观遗传形式，它依据基因来自父本或母本的不同来调节基因的表达。这样，对二倍体的一个印记位点，来源于父本或母本等位基因的表达是不一样的。在每一代中，亲本来源的基因组印记都必须被抹除、重置并保持。这使得印记的基因座易受以上过程中引入的错误所影响。这些错误，加之 DNA 甲基化酶、甲基化 DNA 结合蛋白、组蛋白修饰相关蛋白的基因突变，都与日益增长的人类表观基因组疾病密切相关（图 23-1）。

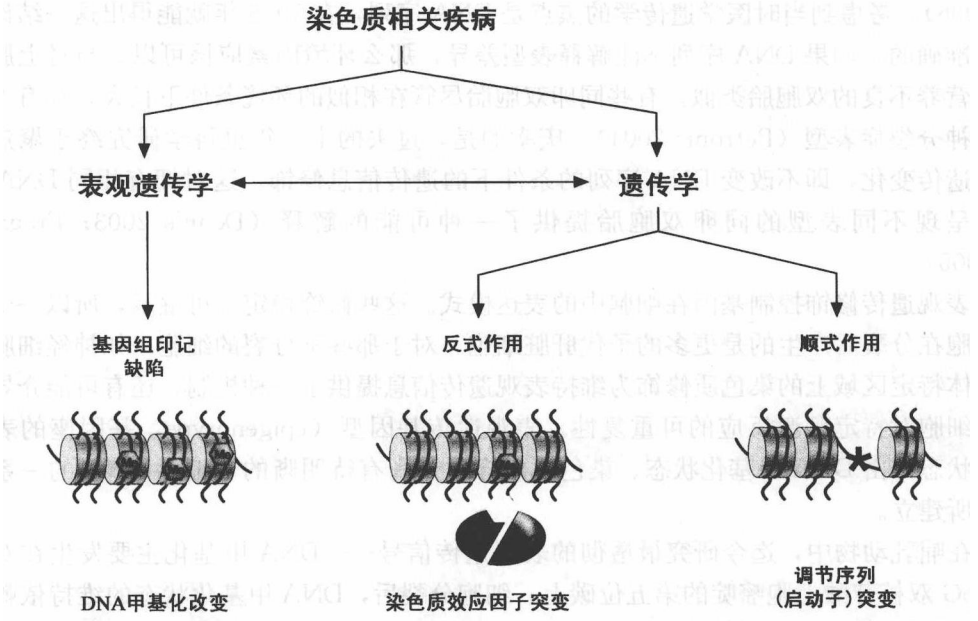



图 23-1 染色质相关疾病的遗传和表观遗传机制

典型的表观遗传机制包括 DNA 甲基化的改变或印记位点染色质的改变，从而破坏单等位基因的表达。遗传机制可以分为两大类。反式效应包括染色质结合因子的缺失或功能性障碍，这些因子可以继而在特定区域改变染色质结构和基因表达。顺式效应指调节所需的非编码区域内发生的突变。这些突变，包括 DNA 重复序列的扩张，可导致染色质改变从而影响基因组的稳定性和基因表达。

## 2. 人类病例研究揭示表观遗传的生物学作用

毫无疑问，模式生物的研究有助于许多生物学原理的理解，尤其是在遗传学、发育生物学和神经生物学领域。然而，我们经常忘记，对于生物学的各个方向，人类本身是最重要的模式生物。对人类数以千计疾病的研究，代表着所有物种中最大规模的突变筛选。如果详尽而系统地进行研究，这些表型在其医学意义以外，还可能会揭示一些生物学原理。所以不必奇怪非孟德尔遗传的“动态突变”（dynamic mutation）是在对脆性 X 染色体综合征患者的研究中得到揭示的（Pieretti et al. 1991）。那些拥有特殊表型的患

者和研究治疗他们的医生常常会开创新的生物学领域并揭示新的遗传分子机制。表观遗传学对人类发育和疾病影响的揭示正是如此。

一位女性患者被她的医生在十年中报道过两次。她因患有膀胱纤维症（CF）、生长激素缺陷所致的矮小于 7 岁时被报道于医学文献（Hubbard et al. 1980）。在寻找 CF 基因的竞争中，Beaudet 等人寻找具有额外表型的 CF 患者，以期发现基因组中小的缺失或染色体重排，加快 CF 基因的定位及鉴定。因此，他们注意到了这个患者。这个女患者当时 16 岁，身高 130cm，拥有正常的智力，但是具有一些明显的身体不对称性（见本章标题右图）。DNA 分析表明她的多个 7 号染色体多态性 DNA 标记，包括着丝粒重复序列为纯合状态（Spence et al. 1988）。排除了非生父和半合子可能性。在分析了患者外祖母（母亲已去世）的 DNA 序列后，Spencer 和其同事发现患者从外祖母身上继承了两个相同的 7 号染色体的着丝粒区域。考虑到 Engel 理论提出人类中可能存在单亲源二倍体（uniparental disomy，UPD）的现象（Engel 1980），Beaudet 和他的同事们立刻认识到母本七号染色体单亲源二倍体携带有 CF 基因隐性突变，并导致了其他的表型。患者的症状及实验室评估，不仅导致了人类单亲源二倍体现象的首次发现，同时也证明父本和母本的基因组是不同的，至少对七号染色体的某些区域是如此。这为我们提供了一种新的非孟德尔遗传原理以解释疾病和发育异常。

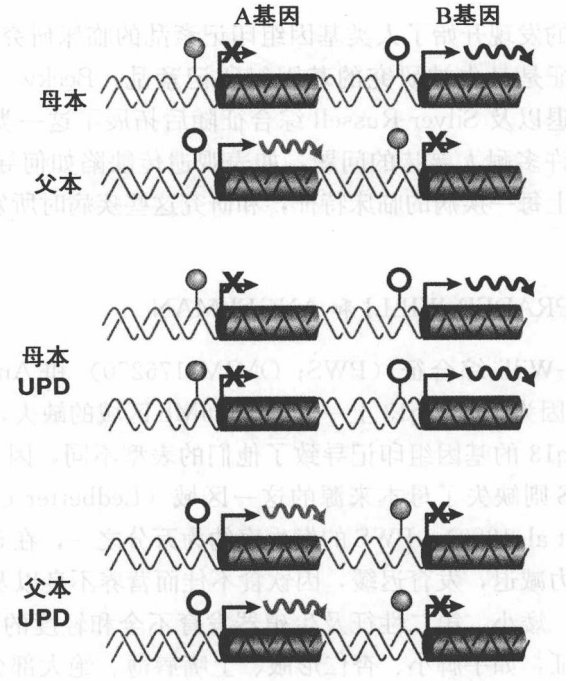



图 23-2　单亲源二倍体（UPD）的结果

实心圆圈指 CpG 岛上游的 DNA 被甲基化，空心圆圈指 CpG 岛上游的 DNA 未被甲基化。DNA 甲基化状态影响其下游基因的表达。母源等位基因倍增（B 基因）；而父源等位基因（A 基因）丢失。

尽管 1988 年，单亲源二倍体现象还被某些人认为是罕见现象，但现在我们已经知道除了 3 号和 19 号染色体外，其他的染色体上都发现了单亲源二倍体现象。

对特殊病例的研究不仅揭示了更多染色体上的单亲源二倍体现象，还在 1989 年提出了单亲源二倍体因改变表观基因型及破坏基因印记而导致疾病的理论（Nicholls et al. 1989）。Nicholls 等研究了一位 Prader-Willi 综合征（PWS）的患者。患者带有 t（13；15）平衡罗伯逊易位（balanced Robertsonian translocation），该易位也存在于其无症状的母亲和其他母系亲属。患者是从母亲处遗传得到第二条 15 号染色体，而所有无症状的后代则从父亲处得到，这使作者们认为母本单亲源二倍体导致了 Prader-Willi 综合征。在确认了第二例具有正常核型的患者也带有 UPD15 后，研究者们提出基因组印记是 Prader-Willi 综合征的病因。进一步，他们推断 15q11-13 的父本缺失或母本单亲源二倍体都会导致 Prader-Willi 综合征，他们也预测 15 号染色体父本单亲源二倍体会导致 Angelman 综合征，正如该区域的母本缺失一样。所有这些推断都已被证实。

## 3. 人类疾病

### 3.1 基因组印记紊乱

单亲源二倍体的发现开始了人类基因组印记紊乱的临床研究。Prader-Willi 综合征和 Angelman 综合征是最先被研究的基因组印记紊乱。Beckwith-wiedemann 综合征、假甲状旁腺机能减退以及 Silver-Russell 综合征随后拓展了这一类疾病的名单，对这些疾病的研究引发了许多耐人寻味的问题，如表观遗传缺陷如何导致疾病表型。在下文中，将简要介绍以上每一疾病的临床特征，和研究这些疾病时所发现的缺陷表型的表观遗传机理。

#### 姐妹综合征：PRADER-WILLI 和 ANGELMAN

大多数 Prader-Willi 综合征（PWS；OMIM 176270）和 Angelman 综合征（AS；OMIM 105830）病因为 15q11-q13 上一个约 5～6Mb 区域的缺失，但他们的表型却完全不同。位于 15q11-q13 的基因组印记导致了他们的表型不同，因为 PWS 缺失了父本来源的该区域，而 AS 则缺失了母本来源的这一区域（Ledbetter et al. 1981；Magenis et al. 1987；Nicholls et al. 1989）。PWS 的发生率约为万分之一，在 50 年前就已经被报道，其特征为婴儿期张力减退，发育迟缓，因饮食不佳而营养不良以及呆滞，随后则出现饮食过量、重度肥胖、矮小、第二性征及生殖器发育不全和轻度的认知障碍。PWS 患者还有特殊的身体特征，如手脚小、杏仁形眼、上嘴唇薄。绝大部分 PWS 患者有轻微或者中度的智力障碍，大部分患者表现出强迫症、易焦虑，有的情绪沮丧孤僻。与 PWS 患者相反，AS 患者具有乐观的情绪，爱笑，有时候会有一阵莫名的大笑。AS 患者具重度发育迟缓、语言能力极差（或完全不具备）、平衡问题（运动失调）、双手不正常的摆动、头小畸形、癫痫以及一些异常的外形，如突出的上颚和宽大的嘴。

张力减退、皮肤及虹膜着色不足和斜视在这两种患者中均有发现。大多数（约

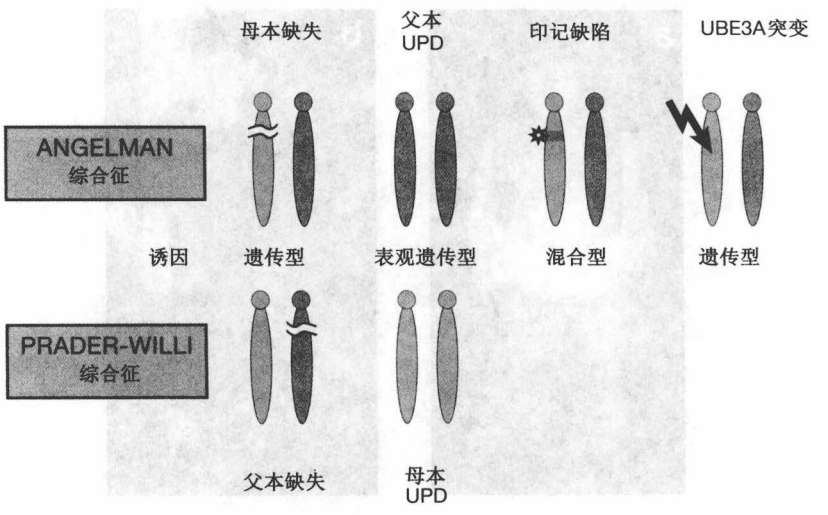



图 23-3　Prader-Willi 综合征和 Angelman 综合征

这两种综合征都可以由遗传缺陷、表观遗传缺陷或两者共同引起。

70%）的 PWS 和 AS 患者的病因分别为父源和母源 15q11-q13 缺失。约 25%的 PWS 患者是由于 15q11-q13 的母本单亲源二倍体，这一区域的父本单亲源二倍体则导致了约 2%～5%的 AS 病例。UPD 在 PWS 和 AS 患者中出现的频率不同通常归因于母源性不分离，因为随女性年龄影响会导致出现 15 号染色体三体或单体。对之随后的“补救”措施导致了母本 UPD（即 PWS 综合征），或父本单亲源二倍体（即 AS 综合征）。这两种 UPD 出现频率的不同可能是因为这两种异常卵子的出现频率和它们被“补救”的频率不同。PWS/AS 重要区域出现易位导致了少于 10%的病例，且这种易位因其亲本来源的性别具很高的家族重现性（近 50%）。事实上，因 15q11-q13 易位或其他异常结构导致在某些家族中同时存在 PWS 和 AS 综合征患者，而表型则由突变的父源或母源性所决定。

**表 23-1　基因印记疾病选列**

| 疾病基因 | 备注 | 涉及的基因 | |
|---|---|---|---|
| Prader-Willi 综合征 | 缺失，UPD，基因印记缺陷 | 15q11-q13 | snoRNA 及其他 |
| Angelman 综合征 | 缺失，UPD，基因印记缺陷，点突变，倍增 | 15q11-q13 | *UBE3A* |
| Beckwith-Wiedemann 综合征 | UPD，基因印记缺陷，11p15.5 倍增，易位，点突变 | 11p15.5 | *IGF2*，*CDKN1C* |
| Silver-Russell 综合征 | UPD，倍增，易位，反转 | 7p11.2 | 区域内若干候选基因 |
| | 表观遗传突变 | 11p15.5 | *H19* 双等位基因表达及 *IGF2* 表达下调 |
| 假性甲状旁腺功能 | 减退症点突变，基因印记缺陷，UPD | 20q13.2 | *GNAS1* |

基因印记缺陷代表另一类导致 PWS 和 AS 综合征的突变。这些缺陷与 15q11-q13

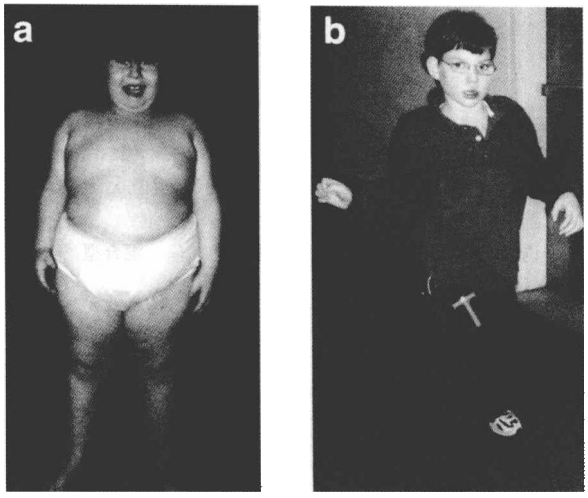



图 23-4 Prader-Willi 综合征患者（a）和 Angelman 综合征患者（b）
同一印记区域的缺陷可导致的综合征有着明显不同的临床特征。图片由 Daniel J. Driscoll 和 Carlos A. Bacino 博士分别提供。

基因印记中心（imprinting center，IC）有关（Ohta et al. 1999），他们导致一条亲本染色体的表观基因型改变，通常变为与相对亲本染色体相同。基因印记缺陷通常是因为基因印记中心（IC）的缺失，但是也有一些病例似乎是因为表观遗传突变造成的，而非 DNA 序列的改变。这些多种多样的基因印记缺陷造成的结果是一样的，包括 DNA 甲基化、染色质结构以及基因表达模式的改变。约 2%～5%的 PWS 和 AS 综合征患者是由于基因印记缺陷造成的，根据缺陷的亲本来源不同，IC 缺失通常有 50%的家族重现性。而无基因印记中心缺失的家族则重现性风险要低。研究发现一些接受过细胞质内精子注射（ICSI）的 AS 患者具有基因印记缺陷，这引出了体外受精易引起基因印记缺陷的可能性（Cox et al. 2002；Orstavik et al. 2003）。未进行 ICSI 而进行了激素治疗的难孕夫妻产出的 AS 综合征患儿也有基因印记缺陷，这提出了新的问题，即是否不孕和基因印记缺陷之间有共同的机制，或辅助生殖技术（激素或 ICSI）具有表观遗传影响（Ludwig et al. 2005）。

AS 综合征中 15q11-q13 受到印记影响的具体基因已经知道。但在 PWS 综合征中还不明了。约 10%～15%的 AS 综合征是由泛素化连接酶 E3（*UBE3A*）基因突变引起，该基因编码 E6 结合蛋白（E6-AP）（Kishino et al. 1997；Matsuura et al. 1997）。通过表达图谱研究，人们发现 *Ube3a* 在小脑蒲肯野氏细胞和海马体神经元细胞中表达，且只从母源染色体表达。进一步的，缺失母源 *Ube3a* 的杂合体小鼠（*Ube3a* +/−）重现了 AS 综合征的特征（Jiang et al. 1998）。这些结果结合人类疾病研究的数据，指出 *UBE3A* 基因是 AS 综合征致病基因。15q11-q13 父本单亲源二倍体或母源性缺失都会导致蒲肯野氏细胞不表达 *UBE3A*。基因印记中心（IC）缺失的病例，似乎失去了抑制性反义转录物的沉默，导致 *UBE3A* 不表达（Rougeulle et al. 1998）。令人费解的是仍有约 10%的 AS 综合征病例的分子机制尚不明了。其中一部分患者的一个染色质重塑蛋白

甲基化 CpG 结合蛋白 2（MBD2）发生了突变。

在 PWS 综合征的患者当中，有若干候选基因只表达于父源等位基因，但还不清楚究竟是哪一个基因导致了 PWS 综合征的表型。目前为止最可能的基因位于非编码的 snoRNA 簇中。最可能的蛋白质编码基因是 *SNURF-SNRPN* 和 *Necdin*（*NDN*）。*SNURF-SNRPN* 基因的主转录起始位点位于基因印记中心，其编码一个调节 RNA 剪切的小 RNA 结合蛋白（SNRPN）。另一基因，“*SNRPN* 上游编码区”（*SNRPN upstream open reading frame*，*SNURF*）及其上游非编码外显子，被认为是基因印记缺陷的主要位点。因为这一基因的敲除会改变 *SNRPN* 基因和其他 15q11-q13 基因的印记。缺失 *Snrpn* 基因的小鼠表型正常，但若同时敲除与其他 15q11-q13 基因同源的某些基因，会导致小鼠减退、发育迟缓并死于断奶之前（Tsai et al. 1999）。若干小核仁 RNA（snoRNA）基因只表达于父源等位基因，也怀疑与 PWS 综合征有关（Meguro et al. 2001）。最近的研究发现如果敲除这些 snoRNA 基因之一（HBII-52）的父源等位基因不会引起 PWS 综合征（Runte et al. 2005）。然而，在小鼠中的研究发现 *Pwcr1/MBII-85* 两个 snoRNA 基因的同时缺失可能是 PWS 模型小鼠中新生鼠致死的原因（Ding et al. 2005）。因此，PWS 综合征可能是由一个或多个 snoRNA 基因缺失引起的，当然也可能同时伴有其他位于 15q11-q13 的父源基因的缺失。对罕见基因易位和缺失家族的细致研究，支持 *PECR1/HBII-85* snoRNA 缺失导致 PWS 综合征的假说（Schule et al. 2005）。

### BECKWITH-WIEDEMANN 综合征

关于 BECKWITH-WIEDEMANN 综合征的故事很好地展示了人类疾病是怎样揭示表观遗传学在正常发育、细胞生长调节以及肿瘤发生中的重要作用。BWS 综合征的主要特征是体细胞过度增长，先天异常，易患小儿胚胎型恶性肿瘤（Weksberg et al. 2003）。BWS 综合征的患者一般身材巨大，舌头肥大，同时伴有偏身肥大，不同程度的耳和其他器官异常，脐突出等特征。同时，很多患者具有肥大的内脏器官、患有胚胎瘤如 Wilm 氏肿瘤、肝母细胞瘤、横纹肌肉瘤等，也会有胰腺增生，经常导致新生儿低血糖。

大多数 BWS 综合征患者是偶发性的，但少数患病家族具有的常染色体显性遗传模式揭示了其病因可能与 11p15 有关（Ping et al. 1989）。BWS 综合征相关肿瘤中母源等位基因的选择性缺失，女性患者对后代的显性遗传，以及某些 BWS 综合征患者具 11p15.5 父源性单亲源二倍体，都表明表观遗传和基因印记必然在 BWS 综合征中起着重要作用，而且 BWS 综合征源于新的或继承而来的一系列遗传及表观遗传的异常。那些与 BWS 综合征相关的印记基因簇大约有 1Mb 大小，分布在 11p15.5，共包含至少 12 个印记基因。这些基因被认为受两个基因印记中心的调控，这两个基因印记中心之间被一个没有基因印记的区域分开（Weksberg et al. 2003）。具有相反基因印记的 *H19* 和类胰岛素生长因子（*IGF2*）以及一个差异性甲基化区域被认为代表了其中的一个基因印记调控区（ICR1）（Joyce et al. 1997；Weksberg et al. 2003）。*H19* 基因编码一个母源性表达的非编码 pol II RNA，*IGF2* 基因则编码一个父源性表达的生长因子。这两个基因

共同拥有一组增强子，对该增强子的接近被基因印记调控区甲基化状态以及一个锌指结合蛋白 CTCF 的结合所影响（Hark et al. 2000）。第二个基因印记调控区（ICR2）含有很多母源性表达的基因，包括周期蛋白依赖性激酶抑制因子（*CDKN1C* 编码的 $p57^{kip2}$），一个钾离子通道组分（KCNQ1），以及一个可能的阳离子转运蛋白（SLC22A1L）。*ICR2* 的差异性甲基化区域位于 *KCNQ1* 的一个内含子上，该区域在父源等位基因上不被甲基化，导致 *KCNQ1* 反义方向上出现 *KCNQ1OT1* 的表达。而母源等位基因 *ICR2* 的甲基化，则被认为抑制了母源性 *KCNQ1OT1* 的表达，从而使母源性基因 *KCNQ1* 和 *CDKN1C* 得以表达（Lee et al. 1999；Smilinich et al. 1999）。

众多表观遗传学和遗传学分子缺陷提供了确认 BWS 综合征致病基因的某些线索。在非甲基化的母源等位基因上，CTCF 结合于 *ICR1*，形成染色质边界从而将 *IGF2* 启动子与其增强子相隔绝。于是这些增强子可以接触到临近染色质边界的 *H19* 启动子，使 *H19* 基因得以转录。父源等位基因 *ICR1* 的甲基化则抑制了 CTCF 的结合，*IGF2* 基因得以表达并抑制 *H19* 基因的表达。包含 *IGF2* 基因座的 11P15.5 的倍增或者该区域的父源性单亲源二倍体患者的发现（均意味着 *IGF2* 基因的过表达），以及过表达 *IGF2* 基因的转基因小鼠出现过度生长以及舌肥大的表型，共同揭示 *IGF2* 基因过表达是导致 BWS 综合征表型的潜在因素之一（Henry et al. 1991；Weksberg et al. 1993；Sun et al. 1997）。有趣的是，与过表达 *IGF2* 基因相似，*CDKN1C* 基因功能缺失突变体也导致 BWS 综合征的发生。缺失 *Cdkn1c* 基因的小鼠具有脐突出现象，但不导致过度生长。然而，在 *Cdkn1c* 基因缺失的同时引入 IGF2 基因的过表达，小鼠则会出现多种 BWS 综合征表型（Caspary et al. 1999）。迄今为止，可引起 BWS 综合征的分子缺陷包括①涵盖 *IGF2* 区域的父源性倍增；②11p15.5 的父源性单亲源二倍体；③*CDKN1C* 母源等位基因的功能缺失；④母源染色体上的易位，这些易位破坏 *KCNQ1*，影响 *IGF2* 基因的印记，有趣的是却不影响 *ICR2*；⑤最常见的则是，*ICR2/KCNQ1OT1* 基因印记的丧失，进而影响 *IGF2* 基因印记，这暗示 *ICR1* 和 *ICR2* 间存在某种调节关系（Cooper et al. 2005）。某些 BWS 综合征患者具有的表观遗传改变，如 *H19 ICR1* 甲基化缺陷，也被发现存在于患有 Wilm 肿瘤却无 BWS 综合征的患者中。这表明表观遗传缺陷发生的时间可能决定了该缺陷是对整个个体还是部分器官发生影响。*ICR1* 甲基化异常经常会导致 Wilms 肿瘤，而 *ICR2* 甲基化异常则易导致 BWS 综合征中的横纹肌肉瘤和肝母细胞瘤。这表明在 11P15.5 存在有多个易诱发肿瘤的基因座（Weksberg et al. 2001；DeBaun et al. 2003；Prawitt et al. 2005）。

### SILVER-RUSSELL 综合征

Silver-Russell 综合征（SRS；OMIM 180860）是一种发育紊乱疾病，主要表现为发育迟缓，身材矮小且不对称，部分具面部、头盖骨以及手指和脚趾畸型。最主要的特征是身体生长不正常，其他特征也变化很大。SRS 有多种遗传病因，但约 10%的患者是由第 7 号染色体的母源性单亲源二倍体所造成（Eggermann et al. 1997）。可能是某父源性表达的促进生长的基因的功能缺失导致了 SRS，但也不能排除是因为某母源性表达的抑制生长的基因出现了过表达。有趣的是，在几例 SRS 患者中发现了由表观遗传突

变引起的染色体 11p15 上 *ICR1* 的去甲基化。这一表观遗传缺陷导致了 *H19* 双等位基因表达和 *IGF2* 表达水平的下降（Gicquel et al. 2005）。

假性甲状旁腺功能减退

在正常副甲状腺激素（PTH）水平下，仍出现甲状旁腺功能减退的各种表型被称为假性甲状旁腺功能减退（PHP）。这些患者对 PTH 具抗性。他们分属多个临床亚型：Ia，Ib，Ic，II 和 Albright 遗传性骨质营养不良（OMIM 103580）。除了甲状旁腺功能减退和骨质营养不良，这些患者同时还可能具有一系列发育及体征缺陷。这些临床上不同的表型源于 *GNAS1* 基因的突变，该基因编码促 α 活性多肽 1（$G_S\alpha$），一种鸟嘌呤核苷结合蛋白。*GNAS1* 位于染色体 20q13.2。*GNAS1* 有三个可变的第一外显子（外显子 *1A*，*XL* 和 *NESP55*），它们和第 2-13 外显子剪接形成不同的转录产物。对于 NESP55 和 XL，这种可变剪接产生不同的蛋白。在这些外显子的周围有着差异性甲基化区域，导致 *NESP55* 只表达于母源等位基因，而 XL、1A 和 NESP55 的一个反义转录物则只表达于父源等位基因。尽管父源和母源的 $G_S\alpha$ 蛋白编码转录物均表达，但母源等位基因在某些组织（如近端肾小管）优先表达。即便对于具有明显常染色体显形遗传模式的突变，基因组印记和组织特异性印记的组合导致了表型和亲源性影响的多样性（Hayward et al. 1998）。值得一提的是一个 *GNAS1* 区域呈父源性单亲源二倍体的患者发展成了 PHP Ib 型疾病（Bastepe et al. 2003）。

通过研究临床病例的基因型与表型，人们发现除 SRS 外的染色体基因印记紊乱（PWS、AS、BWS 和 PHP）均可起源于一系列的遗传学或表观遗传学异常，无论该异常是新生的还是继承的。很难相信这样一个混合的遗传模型只会存在于非常有限的疾病种类中。十多年前，单亲源二倍体只是一种理论上存在的可能性，但现在在许多染色体区域都发现了这种现象并证明其导致了多种疾病和发育表型。人类遗传研究所面临的一个挑战是找到哪些基因导致了与 UPD 相关的表型，以确定哪些疾病可能是由遗传和表观遗传混合机制所导致。

## 3.2　反式作用于染色质结构的紊乱

越来越多的疾病被发现起源于与染色质结构和重塑相关蛋白的突变，这凸显了精密协调的染色质结构对于人类健康的重要性。这些疾病本身没有表观遗传突变，却改变了表观遗传型的决定性组分-染色质状态。这些巨大的表型差异，以及蛋白质水平上细微的改变乃至保守性氨基酸的替换都可能导致人类疾病的现象，正开始为染色质重塑蛋白的严格调控和相互作用提供线索。反式作用于染色质结构的紊乱或者源于染色质重塑蛋白的功能丧失，如 CREB 结合蛋白（CBP），EP300 或者甲基化 CpG 结合蛋白（MeCP2），或者源于 DNA 甲基化相关蛋白的功能丧失，如 DNA 甲基化酶 3B（DNMT3B）或者次甲基四氢叶酸还原酶（MTHFR）（表 23-2）。以上基因中任一基因的功能丧失都导致复杂的多系统的表型或者肿瘤发生，因为它错误地调控大量下游基因的表达。尽管尚未经证实，但是很有可能某些疾病归因于反式作用的非编码 RNA 的突变。

### RUBINSTEIN-TAYBI 综合征

Rubinstein-Taybi 综合征（RSTS，OMIM 180849）的特征包括智力迟钝、拇指和脚趾肥大、面部畸形、先天心脏疾病以及高肿瘤发生率。同卵双胞胎发病的统一性及少量的母源性遗传病例，都暗示了 Rubinstein-Taybi 综合征具有遗传学基础并很可能是常染色体显性遗传。在若干 RSTS 患者中发现了 16p13.3 的细胞学异常（Tommerup et al. 1992）并被定位到含有 CREB 结合蛋白基因（*CREBBP* 或 *CBP*）的区域。*CREBBP* 杂合突变体表明 CBP 的单倍体不足导致 RSTS（Petrij et al. 1995）。CBP 最初报道为 cAMP-应激性结合蛋白 CREB 的一个共激活因子。细胞内 cAMP 水平增加时，蛋白激酶 A（PKA）进入细胞核并磷酸化 CREB，从而将其激活并结合到 cAMP 应激元件上（CRE）（Mayr and Motminy 2001）。CBP 是一个很大的蛋白（约 250 kDa），具有与被 PKA 磷酸化的 CREB 相结合的 bromo 结构域（Chrivia et al. 1993）。CBP 继而通过乙酰化相邻核小体的所有四个核心组蛋白来激活含 CRE 的启动子的转录（Ogryzko et al. 1996）。CBP 还通过其 C 端的某个区域与基本转录因子 TFIIB 直接相互作用（Arias et al. 1994；Kwok et al. 1994）。体外分析一个引起 RSTS 的 CBP 错义突变体（精氨酸 1378 突变为脯氨酸），发现这一突变体失去了 CBP 的组蛋白乙酰转移酶（HAT）活性（Murata et al. 2001）。以上数据以及在小鼠中发现 CBP 单倍体不足导致小鼠学习和记忆能力下降，神经细胞突触可塑性改变以及组蛋白乙酰化异常，都支持 CBP 乙酰转移酶活性的降低是导致 RSTS 表型的重要原因（Alarcon et al. 2004）。与乙酰转移酶活性降低在疾病中所起作用一致的是，最近的研究发现第二个基因 p300 编码一个高效的乙酰转移酶/转录共激活因子，其突变也会引起 RSTS（Roelfsema et al. 2005）。$CBP^{+/-}$ 小鼠中的某些神经细胞突触缺陷以及学习和记忆缺陷，可被组蛋白去乙酰化酶抑制剂恢复（Alarcon et al. 2004），这提出了使用此类药物对 RSTS 综合征患者的智力缺陷进行治疗的可能性。

**表 23-2 反式影响染色质结构的遗传疾病选列**

| 疾病 | 基因 | 备注 |
|---|---|---|
| Rubinstein-Taybi 综合征 | *CREBBP*, *EP300* | |
| Rett 综合征 | *MECP2* | 功能缺失或倍增，引起广泛的表型 |
| α-地中海贫血症和 X 染色质连锁的智力发育迟缓 | *ATRX* | 体细胞突变引起 α-地中海贫血症和骨髓增生异常综合征 |
| ICF 综合征 | *DNMT3B* | |
| Schimke 骨髓免疫发育异常 | *SMARCAL1* | |
| 智力发育迟缓 | *MTHFR* | |

### RETT 综合征

Rett 综合征（RTT，OMIM 312750）是显性的 X 染色体连锁的出生后神经发育紊乱，其主要特征是行动不正常、不协调、癫痫、无意识的双手缠绕以及语言能力下降

(Hagberg et al. 1983)。RTT 综合征被划分为 DSMIV 中类似自闭综合征（ASD）的一类，它与 ASD 有三个主要的共同特征：两者都发生于出生后，通常出现于一段看似正常的发育阶段之后；社交及语言能力的发展被打断；同时均伴有手及手臂的异常特征性移动方式（图 23-5a）。尽管大多数情况下（>99%）RTT 综合征为偶发性，但对少数患病家庭的研究发现 RTT 综合征通过母系遗传，表明 RTT 综合征具有遗传学基础。对这些家族的研究，结合已有的发现即 RTT 综合征主要发生于女性且某些女性携带者没有表型，提出了 RTT 综合征是 X 染色体连锁的显性遗传疾病的假说。使用排除法定位的策略，发现 RTT 基因位于 Xq27-qter，对候选基因的分析表明编码甲基化 CpG 结合蛋白 2（MECP2）的基因是致病基因（Amir et al.，1999）。

MECP2 基因突变是导致 RTT 综合征主要原因的发现，为 RTT 综合征和自闭综合征间的联系提供了分子水平证据。目前已知 *MECP2* 突变在女性中可以引起多种表型，包括学习障碍、孤僻型智力发育迟缓、类 Angelman 综合征以及 ASD 综合征。X 染色体失活（XCI）模式是导致临床表型多样性的决定性分子因素。除少数下效等位基因外，携带 *MECP2* 突变及平衡性 XCI 模式的女性具有典型的 RTT 综合征。而具有偏向野生型等位基因的不平衡性 XCI 模式的女性通常表型较轻（Wan et al. 1999；Carney et al.，2003）。携带 MECP2 突变的男性比女性具有更加广泛的表型，因为男性中该基因座处于半合子状态。导致 RTT 综合征的突变通常引起新生儿死亡，除非该男性只是这种突变体的嵌合体或者具有 XXY 的核型，对于后者，女性患者具有的所有表型也均有发现（Zeev et al. 2000；Neul and Zoghbi 2004）。此外，在女性中基本不导致表型的下效等位基因，在男性中可以导致多重表型，包括智力迟缓、癫痫、颤抖、睾丸增大、双极病或精神分裂症（Meloni et al. 2000；Couvert et al. 2001）。

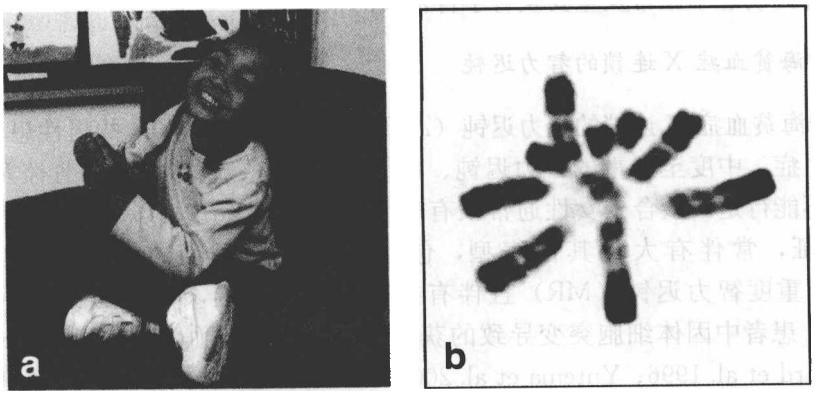



图 23-5　顺式影响染色质的遗传疾病

(a) 这张 Rett 综合征患者的照片展示了异常且固定的手掌移动模式、磨牙以及异常的姿势。照片由 Daniel G. Glaze 博士提供。(b) 来自 ICF 患者的染色体显微照片。照片由 Timothy H. Bestor，Robert A. Rollins and Deborah Bourc'his 博士提供。

MeCP2 因其结合对称性甲基化 CpG 双核苷酸的能力而被发现（Lewis et al. 1992）。它位于异染色质上并作为一种甲基化依赖性转录抑制因子（Nan et al. 1997）。MeCP2

通过其甲基化 CpG 结合结构域与甲基化 DNA 相结合并通过其转录抑制结构域与共抑制因子 Sin3A 和 HDAC 相结合。MeCP2 还与染色质重塑复合物 SWI-SNF 的一个称为 Brahma 的亚基相结合（Harikrishnan et al. 2005）。

RTT 综合征一个耐人寻味的特征是其表型在出生后逐渐出现，却没有神经退行性病变的迹象。通过对 MeCP2 分布及含量的研究，发现其存在于成熟后的神经元中，并可能出现于突触形成后（Shahbazian et al. 2002a；Kishi and Macklis 2004；Mullaney et al. 2004）。这样的分布暗示 MeCP2 在神经成熟并建立其活性后，对神经元功能至关重要，起着调节神经元活性的作用。MeCP2 的部分靶基因已初步得到鉴定，但这些基因中究竟是哪些导致了 RTT 综合征还有待进一步研究（Chen et al. 2003；Martinowich et al. 2003；Horike et al.，2005；Nuber et al. 2005）。对 RTT 综合征患者细胞提取物和丧失了功能性 MeCP2 的小鼠模型的脑提取物进行研究，发现组蛋白乙酰化水平发生了改变（Wan et al. 2001；Shahbazian et al. 2002b；Kaufmann et al. 2005）。这符合根据 MeCP2 与 HDAC 的结合特性所提出的 MeCP2 可能参与组蛋白去乙酰化的推测。有趣的是在小鼠或人中 MeCP2 含量的加倍会导致比 MeCP2 功能缺失更加严重的表型（Collins et al. 2004；Meins et al. 2005；Van Esch et al. 2005）。过表达 MeCP2 是否会对其结合因子产生竞争效应和（或）影响其靶基因的表达还有待明确。为了发现 MeCP2 的潜在新功能，Young 和他的同事发现 MeCP2 可以与 Y 框结合蛋白 1（YB-1）相互作用，该蛋白是一种影响剪切的 RNA 结合蛋白（Young et al. 2005）。MeCP2 调节小报告基因的 RNA 剪切，更重要的是 MeCP2 可能在体内影响 RNA 剪切，因为在 RTT 综合征模型小鼠的脑组织中发生了 RNA 剪切模式的改变（Young et al. 2005）。MeCP2 对神经元特异性基因的表观遗传调节以及 RNA 剪切的重要作用，可能是 RTT 综合征及相关疾病中神经功能不正常及重要发育事件消失的根源。

### α 地中海贫血症 X 连锁的智力迟钝

α 地中海贫血症 X 连锁的智力迟钝（ATRX；OMIM 301040）的男性患者表现为 α 地中海贫血症、中度至严重的智力迟钝、面部特征异常、畸形小头、骨骼及生殖器异常，通常不能行走。杂合子女性通常没有表型。定位于 Xq13 的 *ATRX* 基因的突变导致该综合征，常伴有大量其他表型，包括各种不同程度的 X 连锁的智力迟钝（XLMR），重度智力迟钝（MR）且伴有抽搐的半身不遂，在骨髓发育不良综合征（ATMDS）患者中因体细胞突变导致的获得性 α 地中海贫血症（Gibbons et al. 1995，2003；Villard et al. 1996；Yntema et al. 2002）。ATRX 蛋白含有一个 PHD 锌指结合区域以及一个 SNF 家族的 DNA 依赖性 ATP 酶。由此，并综合考虑到 ARTX 一般位于着丝粒周边异染色质区域并且与异染色质蛋白 1α（HP1α）相结合（McDowell et al. 1999），暗示了它很可能是一个染色质重塑蛋白。ATRX 基因突变引起 α 球蛋白表达量下调以及一些高度重复序列（包括亚端粒区重复序列，Y 染色体特有的卫星序列以及大量核糖体 DNA 序列）的甲基化异常。最近的研究表明 ATRX 基因对于大脑皮层神经元的存活是必须的，这暗示了持续增加的神经元丢失可能导致了 ARTX 突变患者的严重智力迟缓以及痉挛现象（Berube et al. 2005）。

有趣的是，ARTX 的水平受到严格调控并且无论增加或者减少 ATRX 都会引起重要的神经发育问题。例如，有的患者表达正常水平 10%～30%的 ATRX，但仍然具有完全的 ATRX 表型，尽管其具有相当多的 ATRX（Picketts et al. 1996）。过多的ATRX 似乎同样糟糕。过量表达 ATRX 的转基因小鼠具神经管发育缺陷，发育迟缓并死于胚胎期。存活下来的小鼠则具颅面发育异常，强迫性抓挠面部及癫痫。其特征类似于 ATRX 功能缺失患者的临床特征，这意味着 ATRX 的水平很可能受到严格调控，以使含有 ATRX 的蛋白复合物具有正常的功能。

### 免疫缺陷、着丝粒区域不稳定以及面部异常综合征

免疫缺陷、着丝粒区域不稳定以及面部异常综合征（ICF，OMIM 242860）是一种罕见的常染色体隐性染色体断裂疾病。ICF 患者具有两种固定的表型——免疫缺陷和细胞学异常。易变及相对不明显的表型包括颅面骨缺陷，如扁平的鼻梁、内眦赘皮、高额头低耳位、精神运动迟缓以及肠道官能障碍（Smeets et al. 1994）。免疫缺陷症状通常很严重，并经常因呼吸道及胃肠道感染导致幼年期夭折。血清中 IgG 含量的下降是最常见的免疫症状，B 细胞和 T 细胞的减少也在某些病例中有报道（Ehrlich 2003）。根据对 ICF 患者血细胞以及培养细胞（图 23-5b）的核型分析结果，发现细胞学异常主要影响 1 号和 16 号染色体，少数情况下 9 号染色体也有异常（Tuck-Muller et al. 2000）。

远在 ICF 综合征被发现之前，就已经发现了 1 号、9 号以及 16 号染色体着丝粒周边重复序列的低甲基化现象（Jeanpierre et al. 1993）。这些染色体具有最大的重复分布于着丝粒附近的典型卫星序列（卫星 2、卫星 3）。DNA 甲基转移酶基因（*DNMT3B*）的功能缺失突变导致 ICF 综合征的发现，为着丝粒周边卫星序列 2 和 3 的低甲基化提供了线索（Hansen et al. 1999；Okano et al. 1999；Xu et al. 1999）。然而，仍不清楚为什么缺失一个广泛表达的 DNA 甲基转移酶的功能会选择性影响特定的重复序列。一个可能的解释是 DNMT3B 和/或其结合蛋白有着特定的亚细胞水平分布（Bachman et al. 2001）。另一个可能的解释是，基因组上高密度的 CpG 重复序列，如卫星序列 2（Gowher and Jeltsch 2000）以及在面肩胛臂肌肉萎缩综合征中 D4Z4 重复序列（Kondo et al. 2000）对 DNMT3B 的活性有更高的要求。是否还有其他特定的序列也发生低甲基化还有待于进一步确定，但可以预见 DNA 低甲基化会改变那些在颅面、神经系统以及免疫发育中起重要作用的基因的表达。

对来源于 ICF 综合征患者的成淋巴细胞瘤细胞系的 RNA 与正常对照进行基因表达研究，发现若干与淋巴细胞成熟、迁移、激活以及归巢相关的基因表达发生了改变（Ehrlich et al. 2001）。然而，不清楚的是为什么 DNMT3B 的缺失可以引起这些基因的失控，因为这些基因启动子的甲基化模式并没有改变。考虑到目前在 ICF 综合征患者当中只发现了卫星 DNA 序列发生了低甲基化，有假说认为某些在 ICF 患者中表达发生改变的基因位于卫星 DNA 序列附近。这些序列被甲基化的时候表现为典型的异染色质，因此在 ICF 患者中，富含卫星序列 2 和 3 的异染色质区域的异常影响了这些基因的表达（Bickmore and van der Maarel 2003）。

SCHIMKE IMMUNO-OSSEUS 综合征

Schimke immuno-osseous 发育异常综合征（SIOD，OMIM 242900）是一种常染色体隐性多系统的综合征，特征为脊柱以及长骨末端发育异常、生长发育缺陷、点状或块状肾小球硬化导致的肾脏功能异常以及 T 细胞介导的免疫缺陷（Schimke et al. 1971；Spranger et al. 1991）。SIOD 综合征是由于 *SMARCAL1*（*SWI/SNF2-related*，*matrix associated*；*actin-dependent regulator of chromatin*，*sub-family a-like1*）基因突变引起的，*SMARCAL1* 编码一个被认为通过染色质重塑调节转录活性的蛋白（Boerkoel et al. 2002）。无意及移码突变会引起严重的表型，而某些错义突变只导致轻微的或者部分表型（Boerkoel et al. 2002）。最近，一个患有 B 细胞淋巴瘤以及 SIOD 综合征的患者被发现在 *SMARCAL1* 基因上有突变，这暗示失去该蛋白的功能可能导致致命的淋巴细胞瘤综合征（Taha et al. 2004）。*SMARCAL1* 缺失导致 SIOD 综合征表型的确切机理还有待于进一步阐明。

次甲基四氢叶酸还原酶缺陷

次甲基四氢叶酸还原酶（MTHFR）参与将 5，10-次甲基四氢叶酸（5，10-MTHF）转化为 5-次甲基四氢叶酸（5-MTHF）的过程。由同型半胱氨酸合成甲硫氨酸所需的甲基来自于 5-MTHF。甲硫氨酸进一步转化为 S-腺苷甲硫氨酸（SAM），所有甲基化转移酶的主要甲基供体。MTHFR 的缺陷引起一种少见的常染色体隐性综合征，主要特征是智力迟缓（Rozen 1996）。一种常见的热敏感多态性（677C⟶T，将丙氨酸突变为缬氨酸）引起 MTHFR 的活性降低，并且在纯合子状态下若食物中缺乏叶酸则常常伴有高同型半胱氨酸血症（Goyette et al. 1994）。该多态性还是动脉硬化，神经管缺陷以及癌症的潜在诱因（Ma et al. 1997；Brattstrom et al. 1998；Chen et al. 1999；Botto and Yang 2000；Schwahn and Rozen 2001）。MTHFR 缺陷的纯合子和杂合子小鼠都具有 SAM 含量及 DNA 甲基化整体水平下降的表型。此外，纯合突变体具有大动脉脂沉积和神经衰退的表型（Chen et al. 2001）。伴随着 MTHFR 部分或完全缺失的 DNA 甲基化整体水平改变，暗示着某些 MTHFR 功能缺失导致的表型可能因 DNA 甲基化水平下降对染色质发生影响而造成。MTHFR 缺陷曾有一例被报道引起 Angelman 综合征（Arn et al. 1998），而且重度 MTHFR 缺陷与 AS 和 RTT 综合征有许多相同的表型（Farral-Valevski et al. 2000）。

## 3.3 顺式作用于染色质结构的紊乱

对于遵循孟德尔规律的疾病，找到其致病基因的方式，通常是发现发生于外显子或剪切位点的突变，从而导致基因产物（RNA 或蛋白）发生改变或者不能产生。然而对于许多疾病，总有一些患者，尽管与该致病基因位点连锁，但对其基因编码区及非编码区测序后发现没有任何突变。顺式影响基因表达的表观遗传或遗传异常，日渐清晰地被认为是某些无外显子突变的孟德尔疾病病例的病因。以下是染色质结构的顺式改变可以导致人类疾病的三个例证（表 23-3）。

表 23-3　顺式影响染色质结构的遗传疾病选列

| 疾病 | 基因 | 备注 |
| --- | --- | --- |
| Aδβ-以及 δβ-地中海贫血症 | LCR 缺失引起珠蛋白表达下降 | |
| 脆性 X 染色体 | 综合征 CCG 重复序列的扩增导致甲基化异常以及 *FMR1* 基因沉默 | 前突变等位基因(60～200)引起神经退行性疾病 |
| FSH 营养不良 | D4Z4 重复序列变短导致抑制态染色质减少 | |
| 多发癌 | *MLH1* 在生殖干细胞内的表观突变 | |

### αδβ 和 δβ 地中海贫血症

地中海贫血症是世界上最普遍的单基因疾病。其病因各不相同，但均因一条或多条珠蛋白链表达下降，导致血红蛋白合成紊乱。各珠蛋白链合成的不平衡导致红血球生成异常及重度贫血（Weatherall et al. 2001）。数以百计的编码和剪切突变体已被发现，但调节序列缺失的发现指出了染色质结构的改变是如何引起某些地中海贫血症亚型的。缺失 β 珠蛋白基因上游约 100kb 的序列（但该基因仍保持完整）可以导致 αδβ 地中海贫血症，由此发现了调控 β 珠蛋白表达的基因座控制区域（locus control region，LCR）（Kioussis et al. 1983；Forrester et al. 1990）。小一些的 LCR 区域序列缺失可以导致 δβ 地中海贫血症（Curtin et al. 1985；Driscoll et al. 1989）。尽管这些缺失序列在 β 珠蛋白基因编码区上游几十 kb 处，但是这些缺失会导致 β 珠蛋白基因座的染色质结构变化（Grosveld 1999）。

### 脆性 X 染色体综合征

脆性 X 染色体智力迟缓（OMIM 309550）是最常见的遗传性智力迟缓病之一。60 多年以前，Martin 和 Bell 报道了一个 X 染色体连锁的智力迟缓家系（Martin and Bell 1943）。在 1969 年，Lubes 在若干智力迟缓的男性患者和一个没有症状的女性上发现 X 染色体长臂缩短了（Lubs 1969）。该染色体异变被定位到 Xq27.3 并被称为脆性 X 染色体（Harrison et al. 1983）。细胞学研究，尤其是使用无叶酸和胸腺嘧啶的培养基进行的研究，揭示了 X 染色体连锁的智力迟缓患者中的脆性位点，随后他们在诊断上称作脆性 X 染色体综合征（Sutherland 1977；Richards et al. 1981）。患病的男性具有中度至重度的智力迟缓，不正常的组织肥大，结缔组织异常，如关节过度延伸，大耳朵（图 23-6）（Hagerman et al. 1984）。导致脆性 X 染色体综合征的基因是 *FMR1*，该基因编码 FMRP 蛋白。导致突变的最常见的机制是不稳定的非编码 CGG 重复序列的扩增（Warren and Sherman 2001）。正常等位基因含 6～60 个重复序列，前突变等位基因含 60～200 个重复序列，而完全突变等位基因具有 200 个以上的重复序列。重复序列在 *FMR1* 基因 5′非翻译区的扩增为染色质结构顺式改变导致基因紊乱提供了非常好的证据。在完全突变的病例中，由于重复序列的扩增导致位于 *FMR1* 基因 5′调控区域的 CpG 岛的异常甲基化（Verkerk et al. 1991）。有报道表明脆性 X 染色体综合征患者中该基因 5′端的组蛋白乙酰化水平较正常人低（Coffee et al. 1999）。相应的，DNA 甲基化和组蛋白

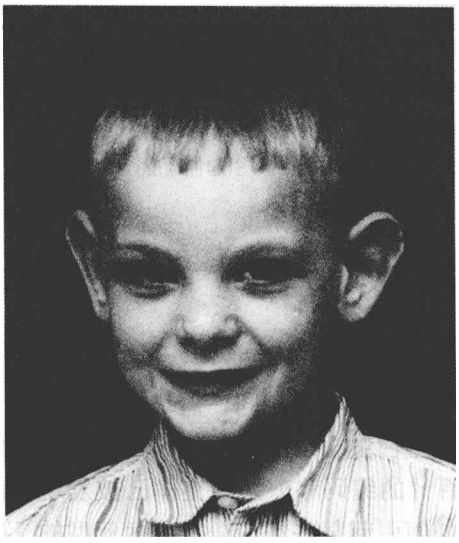


图 23-6 反式影响染色质的遗传疾病的例子 这是一个脆性 X 染色体综合征患者的照片，除智力低下外，典型特征为显著的前额和大耳朵。图片由 Stephen T. Warren 博士提供。

乙酰化模式的改变导致了 *FMR1* 基因失去表达，即 FMRP 在脆性 X 染色体综合征患者中失去功能。这样，这些患者中有原发性基因突变及继发性表观遗传突变。

一个有趣的表观遗传机制被用于解释 *FMR1* 基因的 CGG 重复序列是如何被甲基化并进而沉默的。前突变 CGG 重复序列形成一个单一稳定的发卡结构（Hanada et al. 2003），且 rCGG 重复序列可以被 Dicer 剪切。基于这两个发现人们提出了一种可能，扩增的 CGG 重复序列（在发育早期未被甲基化）可以转录，产生的 RNA 形成发卡结构被 Dicer 剪切产生小非编码 RNA。这些小 RNA 分子与 RNA 诱导的转录沉默起始因子（RITS）相结合并募集从头 DNA 甲基化酶和组蛋白甲基化酶到 *FMR1* 基因的 5′非编码区，导致了 CGG 重复序列的完全甲基化以及在发育过程中 *FMR1* 基因的转录逐渐被抑制（Jin et al. 2004a）。

FMRP 是特异性 RNA 结合蛋白，它含有 2 个 KH 结构域和一个 RGG 盒。FMRP 以 RNA 依赖的方式通过信使核蛋白部分与多聚核糖体相结合，在体内和体外均抑制翻译（Laggerbauer et al. 2001；Li et al. 2001）。结合了 mRNA 和多聚核糖体的 FMRP 位于树突脊，这为其通过调控特定的蛋白合成来应答突触的刺激提供了证据（Feng et al. 1997；Weiler and Greenough 1999；Brown et al. 2001；Darnell et al. 2001，2005）。FMRP 的潜在靶基因已被发现，并被证实在突触发育中起一定作用，且可以部分解释患者的神经发育表型（Brown et al. 2001；Darnell et al. 2001）。

若干研究认为 RNA 干扰途径是 FMRP 调控翻译的主要机制。果蝇脆性 X 同源基因（*Dfmr1*）与 Argonaute（*AGO2*）和 RNA 诱导的沉默复合物（RISC）相结合，哺乳动物 FMRP 与 $EIF_2C2$ 相互作用并且结合有 Dicer 活性（Caudy et al. 2002；Ishizuka et al. 2002；Jin et al. 2004b）。最流行的解释 FMRP 作为翻译抑制因子的假说机制是 FMRP 结合到特殊的 RNA 配体上，连同 miRNA 一起募集 RISC，促进 miRNA 和 mRNA配体间的识别（Jin et al. 2004a）。

脆性 X 染色体前突变（60～200 重复）携带者患有一种特殊的神经退行性疾病，其特征为战栗和运动性共济失调（Hagerman and Hagerman 2004）。有趣的是，前突变体可能会在 RNA 水平上导致疾病，因为他们具有 FMR1 的 RNA 和蛋白，。在动物模型中研究发现 CGG 重复序列编码的 RNA 会与一些细胞内蛋白结合并且改变它们的功能，引起蛋白的聚集（Jin et al. 2003；Willemsen et al. 2003）。

面肩胛臂营养失调

面肩胛臂营养失调（FSHD；OMIM 158900）是一种常染色体显性肌肉失调，其特征为面部，上臂以及肩膀的肌肉逐渐萎缩。严重的患者会失聪，极少部分的重度患儿智力发育迟缓并具有癫痫（Mathews 2003）。*FSHD*（*FSHD1*）基因的主要基因座在靠近 D4Z4 的 4q35 的亚端粒区域内，这是一个富含 GC 的 3.3Kb 低拷贝重复序列。这些重复序列具多态性且在正常染色体上有 11～150 个单位，而在 FSHD 染色体上为 1～10 个单位（Wijmenga et al. 1992；van Deutekom et al. 1993）。D4Z4 末端的另一种可变卫星重复序列（β-68bp *Sau3A*）在 FSHD 综合征发生中似乎起着一定的作用。β 卫星重复序列的 4qA 变体，以及 D4Z4 重复序列的变短，为 FSDH 综合征所必须（Lemmers et al. 2002）。D4Z4 的变短和 4qA β 卫星重复序列究竟如何导致该疾病还不清楚。含有 D4Z4 的 4q35 区域与其他亚端粒相似，具有典型的异染色质特征（Flint et al. 1997；Tupler and Gabellini 2004）。体内和体外的研究发现 D4Z4 内的一个长为 27bp 的序列可以和 D4Z4 抑制复合物（DRC）相结合，DRC 复合物包括转录抑制因子 YY1，高泳动性蛋白 2（HMGB2）以及核仁蛋白（Gabellini et al. 2002）。Bickmore，van der Maarel，Gabellini 和他们的同事推测由于重复序列的变短导致染色质状态的抑制性降低，从而造成 4q35-q 区域的基因表达上升（Bickmore and van der Maarel 2003；Tupler and Gabellini 2004）。与该假说吻合的是 *FRG1*、*FRG2*（FSHD 区域基因 1 和 2）以及腺嘌呤核苷转运子 1（*ANT1*）在 FSHD 综合征患者的肌细胞中高表达（Gabellini et al. 2002）。当然 D4Z4 重复序列的变短是直接还是间接导致这些基因的表达水平变化以及这些基因对于疾病的表型有什么作用还有待于进一步研究。

表观突变和人类疾病

在两个患有多种癌症的患者身上发现了 DNA 错配修复基因 *MLH1* 的表观突变（Suter et al. 2004）。在这两个患者所有可以检测的组织上都发现了 *MLH1* 基因启动子的异常甲基化。在肿瘤组织中，敲除或者丢失 *MLH1* 杂合性都会导致 MLH1 蛋白的完全消失。两个患者都患有直肠癌；其中一个患有十二指肠癌另外一个患有子宫内膜和乳腺癌以及胎记瘤。只有当科研人员开始系统性寻找表观遗传突变时，其在人类疾病中的作用才会越来越清晰。

## 3.4　表观遗传与环境的相互作用

来自于人类和动物模型的研究数据表明，环境可以影响表观遗传标记从而影响基因功能。同卵双胞胎在出生早期具有相似的表观遗传型，但随后却在 5-甲基胞嘧啶和组蛋白乙酰化的分布和密度上产生了明显的不同。这有力地证明了表观遗传型是亚稳态的并具时间多样性（Fraga et al. 2005）。很可能多种环境因素和随机事件的参与导致了表观基因组的多样性（图 23-7）（Anway et al. 2005），但食物和早期经历似乎是决定性因素。

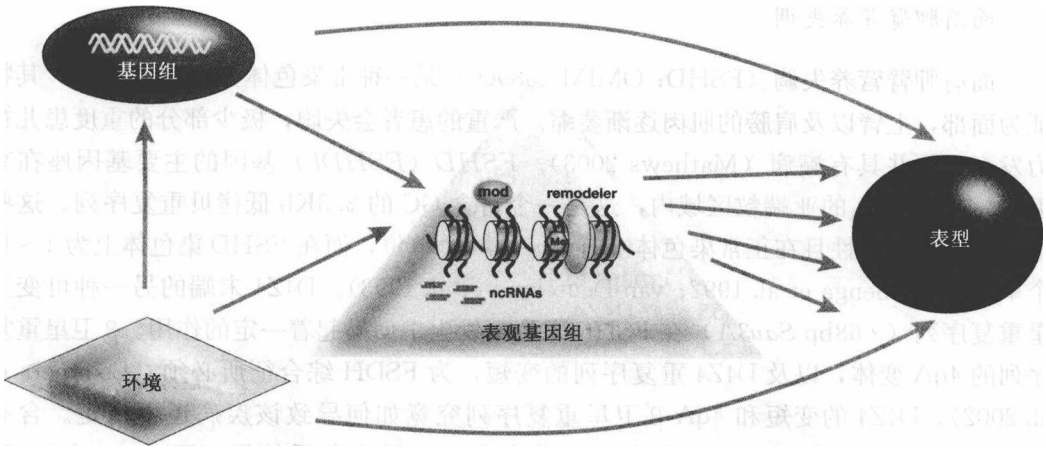



图 23-7 表观遗传型在和基因型以及环境因素一起共同决定表型时起到了至关重要的作用。已知影响基因表达和基因组稳定性的表观遗传因素包括 DNA 甲基化、染色质重塑复合物、组蛋白共价修饰、组蛋白变体和非编码调节 RNA（ncRNA）。

### 食物以及表观遗传型在衰老和疾病中的作用

许多报道都表明随着年龄的增加，整体的 DNA 甲基化水平会下降，与此同时有些特殊位点也可能高甲基化（Hoal-van Heden and van Helden 1989；Cooney 1993；Rampersaud et al. 2000）。考虑到大量的数据将 DNA 甲基化的改变和癌症的风险或者进程相关，表观遗传变化可能会随着年龄增长而提高患癌的风险。在患有尿毒症并且接受血液透析的成年男性患者中，食物对于控制整体甲基化水平的作用已经被很好地证明了。这些患者血液中的高同型半胱氨酸暗示了低甲硫氨酸可能是因叶酸缺乏所致。他们的整体或一些特殊位点的 DNA 甲基化水平都会下降，但经高剂量的叶酸治疗后 DNA 甲基化水平会恢复（Ingrosso et al. 2003）。

因为由缺乏叶酸及维生素 B12 导致的许多神经心理特征与一些散发性神经心理疾病的特征相同，由此推测后者可能是由于中央神经系统甲基化模式变化而引起的（Reyolds et al. 1984）。在叶酸有效的忧郁症患者中发现低水平的 SAM；进一步的研究表明，添加 SAM 对于某些忧郁症患者作为辅助治疗是有效的（Bottiglieri et al. 1994）。最终，尽管还不知道孕妇增加叶酸摄入是如何降低胎儿神经管缺陷的概率，但是非常可能的是影响了表观遗传介导的 DNA 和组蛋白甲基化。在食物中添加叶酸，维生素 B12 以及甜菜碱会改变 *agouti viable yellow* 小鼠子代的表观遗传型和表型，这也许是尚有待发现的众多人类和哺乳动物中例证中的首例（Wolff et al. 1998；Waterland and Jirtle 2003）。

### 早期经历和表观遗传型

目前为止，最好的关于早期经历以及母亲行为改变哺乳动物的表观遗传型的例子只

在大鼠中报道过。大鼠母亲经常的舔舐疏理幼鼠会导致幼鼠海马体内 GR 基因启动子区域甲基化状态的改变。相比那些很少被舔舐疏理的幼鼠，那些经常被舔舐疏理的幼鼠在 GR 基因的启动子区域 DNA 甲基化水平会下降，组蛋白乙酰化水平会上升（Weaver et al. 2004）。由于表观遗传型的改变导致 GR 水平的上升影响了幼鼠应激激素水平的调控以及终生的对于压力的反应（Liu et al. 1997；Weaver et al. 2004）。尽管在人类中还没有类似的例证，但是越来越多的疑问被提出，那就是在人类中早期经历对于调节表观遗传基因型和导致精神疾病到底起什么样的作用。

## 4. 展望未来

在未来十年，我们预计导致表观遗传基因型改变的突变将逐渐地被认识，它们是引起各种人类疾病的突变机制。传统上，鉴定致病基因的主要手段是研究携带染色体异常的家系，进行基因定位克隆。当前，当我们试图去发现某些更常见更具破坏性的疾病（如精神分裂症、自闭症以及情绪紊乱）的病因时，我们面临严峻的挑战。因为这些疾病的家系并不常见，遗传的异质性可能很高。而遗传学数据尤其是同卵双胞胎不一致的比例并不总支持简单明了的孟德尔遗传模型。这些发现，再结合环境因素对于某些疾病的深远影响，都强调了在这些疾病中研究表观遗传型的重要性。甚至单基因疾病，如 Angelman 综合征、Beckwith-Wiedemann 综合征以及 Silver-Russell 综合征都可以因基因组突变或因表观遗传型突变而引起，这些突变既可以是遗传的也可以是新形成的。这类分子水平的变化毫无疑问也将在其他人类疾病中出现。更进一步来说，研究数据表明许多涉及表观遗传调控的蛋白都受到严格调控，因此缺失突变或者基因倍增都会引起人类疾病，同时还暗示着影响转录、RNA 剪接以及蛋白修饰的表观遗传突变也会引起各种疾病。

## 致　谢

我们感谢 Timothy H. Bestor 博士、Robert A. Rollinsyi 博士以及 Deborah Bourc'his 博士为我们提供来源于 ICF 综合征患者的染色体图片；Daniel J. Driscoll 博士为我们提供 Prader-Willi 综合征患者的图片，Carlos A. Bacino 博士和 Daniel G. Glaze 博士提供 Rett 综合征患者的图片以及 Stephen T. Warren 博士的脆性 X 染色体综合征患者的图片。我们非常感谢 Zoghbi 实验室的同事阅读本章并提出了很好的建议。我们也要感谢过去和现在致力于 Rett 综合征、Prader-Willi 综合征以及 Angelman 综合征研究的实验室成员们。最后，同样重要的，我们感谢那些患有 Rett 综合征、7 号染色体单亲源二倍体、Prader-Willi 综合征、Angelman 综合征，还有孤独症患者以及他们的家人，是他们启发了我们，让我们明白表观遗传在人类疾病中的作用。我们的资金来源于以下机构：国立健康研究所（5 P01 HD040301-05；5 P30 HD024064-17；5 P01 HD37283）；国际 Rett 综合征协会；Rett 综合征研究基金会；立即治愈孤独症基金会；Simons 基金会；March of Dimes 基金会（12-FY03-43）以及 Blue Bird 诊所 Rett 中心。H. Y. Z 是 How-

ard Hughes 医学院的研究员。由于篇幅原因很多重要的相关文献未被列出，我们深表遗憾。

袁 文 译
谢精卫 校
朱 冰 审校

## 参考文献

Alarcon J.M., Malleret G., Touzani K., Vronskaya S., Ishii S., Kandel E.R., and Barco A. 2004. Chromatin acetylation, memory, and LTP are impaired in CBP$^{+/-}$ mice: A model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. *Neuron* **42:** 947–959.

Amir R.E., Van den Veyver I.B., Wan M., Tran C.Q., Francke U., and Zoghbi H.Y. 1999. Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23:** 185–188.

Anway M.D., Cupp A.S., Uzumcu M., and Skinner M.K. 2005. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* **308:** 1466–1469.

Arias J., Alberts A.S., Brindle P., Claret F.X., Smeal T., Karin M., Feramisco J., and Montminy M. 1994. Activation of cAMP and mitogen responsive genes relies on a common nuclear factor. *Nature* **370:** 226–229.

Arn P.H., Williams C.A., Zori R.T., Driscoll D.J., and Rosenblatt D.S. 1998. Methylenetetrahydrofolate reductase deficiency in a patient with phenotypic findings of Angelman syndrome. *Am. J. Med. Genet.* **77:** 198–200.

Bachman K.E., Rountree M.R., and Baylin S.B. 2001. Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *J. Biol. Chem.* **276:** 32282–32287.

Bastepe M., Frohlich L.F., Hendy G.N., Indridason O.S., Josse R.G., Koshiyama H., Korkko J., Nakamoto J.M., Rosenbloom A.L., Slyper A.H., et al. 2003. Autosomal dominant pseudohypoparathyroidism type Ib is associated with a heterozygous microdeletion that likely disrupts a putative imprinting control element of *GNAS*. *J. Clin. Invest.* **112:** 1255–1263.

Berube N.G., Mangelsdorf M., Jagla M., Vanderluit J., Garrick D., Gibbons R.J., Higgs D.R., Slack R.S., and Picketts D.J. 2005. The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. *J. Clin. Invest.* **115:** 258–267.

Bickmore W.A. and van der Maarel S.M. 2003. Perturbations of chromatin structure in human genetic disease: Recent advances. *Hum. Mol. Genet.* **12:** R207–R213.

Boerkoel C.F., Takashima H., John J., Yan J., Stankiewicz P., Rosenbarker L., Andre J.L., Bogdanovic R., Burguet A., Cockfield S., et al. 2002. Mutant chromatin remodeling protein SMARCAL1 causes Schimke immuno-osseous dysplasia. *Nat. Genet.* **30:** 215–220.

Bottiglieri T., Hyland K., and Reynolds E.H. 1994. The clinical potential of ademethionine (S-adenosylmethionine) in neurological disorders. *Drugs* **48:** 137–152.

Botto L.D. and Yang Q. 2000. 5,10-Methylenetetrahydrofolate reductase gene variants and congenital anomalies: A HuGE review. *Am. J. Epidemiol.* **151:** 862–877.

Brattstrom L., Wilcken D.E., Ohrvik J., and Brudin L. 1998. Common methylenetetrahydrofolate reductase gene mutation leads to hyperhomocysteinemia but not to vascular disease: The result of a meta-analysis. *Circulation* **98:** 2520–2526.

Brown V., Jin P., Ceman S., Darnell J.C., O'Donnell W.T., Tenenbaum S.A., Jin X., Feng Y., Wilkinson K.D., Keene J.D., et al. 2001. Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107:** 477–487.

Carney R.M., Wolpert C.M., Ravan S.A., Shahbazian M., Ashley-Koch A., Cuccaro M.L., Vance J.M., and Pericak-Vance M.A. 2003. Identification of MeCP2 mutations in a series of females with autistic disorder. *Pediatr. Neurol.* **28:** 205–211.

Caspary T., Cleary M.A., Perlman E.J., Zhang P., Elledge S.J., and Tilghman S.M. 1999. Oppositely imprinted genes *p57$^{Kip2}$* and *Igf2* interact in a mouse model for Beckwith-Wiedemann syndrome. *Genes Dev.* **13:** 3115–3124.

Caudy A.A., Myers M., Hannon G.J., and Hammond S.M. 2002. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev.* **16:** 2491–2496.

Chen J., Giovannucci E.L., and Hunter D.J. 1999. MTHFR polymorphism, methyl-replete diets and the risk of colorectal carcinoma and adenoma among U.S. men and women: An example of gene-environment interactions in colorectal tumorigenesis. *J. Nutr.* **129:** S560–S564.

Chen W.G., Chang Q., Lin Y., Meissner A., West A.E., Griffith E.C., Jaenisch R., and Greenberg M.E. 2003. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302:** 885–889.

Chen Z., Karaplis A.C., Ackerman S.L., Pogribny I.P., Melnyk S., Lussier-Cacan S., Chen M.F., Pai A., John S.W., Smith R.S., et al. 2001. Mice deficient in methylenetetrahydrofolate reductase exhibit hyperhomocysteinemia and decreased methylation capacity, with neuropathology and aortic lipid deposition. *Hum. Mol. Genet.* **10:** 433–443.

Chrivia J.C., Kwok R.P., Lamb N., Hagiwara M., Montminy M.R., and Goodman R.H. 1993. Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* **365:** 855–859.

Coffee B., Zhang F., Warren S.T., and Reines D. 1999. Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells (erratum *Nat. Genet.* **22:** 209 [1999]). *Nat. Genet.* **22:** 98–101.

Collins A.L., Levenson J.M., Vilaythong A.P., Richman R.D., Armstrong L., Noebels J.L., Sweatt J.D., and Zoghbi H.Y. 2004. Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.* **13:** 2679–2689.

Cooney C.A. 1993. Are somatic cells inherently deficient in methylation metabolism? A proposed mechanism for DNA methylation loss, senescence and aging. *Growth Dev. Aging* **57:** 261–273.

Cooper W.N., Luharia A., Evans G.A., Raza H., Haire A.C., Grundy R., Bowdin S.C., Riccio A., Sebastio G., Bliek J., et al. 2005. Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* **13:** 1025–1032.

Couvert P., Bienvenu T., Aquaviva C., Poirier K., Moraine C., Gendrot C., Verloes A., Andres C., Le Fevre A.C., Souville I., et al. 2001. *MECP2* is highly mutated in X-linked mental retardation. *Hum. Mol. Genet.* **10:** 941–946.

Cox G.F., Burger J., Lip V., Mau U.A., Sperling K., Wu B.L., and Horsthemke B. 2002. Intracytoplasmic sperm injection may increase the risk of imprinting defects. *Am. J. Hum. Genet.* **71:** 162–164.

Curtin P., Pirastu M., Kan Y.W., Gobert-Jones J.A., Stephens A.D., and Lehmann H. 1985. A distant gene deletion affects β-globin gene function in an atypical γζβ-thalassemia. *J. Clin. Invest.* **76:**

1554–1558.

Darnell J.C., Jensen K.B., Jin P., Brown V., Warren S.T., and Darnell R.B. 2001. Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* **107:** 489–499.

Darnell J.C., Fraser C.E., Mostovetsky O., Stefani G., Jones T.A., Eddy S.R., and Darnell R.B. 2005. Kissing complex RNAs mediate interaction between the Fragile-X mental retardation protein KH2 domain and brain polyribosomes. *Genes Dev.* **19:** 903–918.

DeBaun M.R., Niemitz E.L., and Feinberg A.P. 2003. Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of *LIT1* and *H19*. *Am. J. Hum. Genet.* **72:** 156–160.

Dennis C. 2003. Epigenetics and disease: Altered states. *Nature* **421:** 686–688.

Ding F., Prints Y., Dhar M.S., Johnson D.K., Garnacho-Montero C., Nicholls R.D., and Francke U. 2005. Lack of *Pwcr1/MBII-85* snoRNA is critical for neonatal lethality in Prader-Willi syndrome mouse models. *Mamm. Genome* **16:** 424–431.

Driscoll M.C., Dobkin C.S., and Alter B.P. 1989. γδβ-thalassemia due to a de novo mutation deleting the 5′ β-globin gene activation-region hypersensitive sites. *Proc. Natl. Acad. Sci.* **86:** 7470–7474.

Eggermann T., Wollmann H.A., Kuner R., Eggermann K., Enders H., Kaiser P., and Ranke M.B. 1997. Molecular studies in 37 Silver-Russell syndrome patients: Frequency and etiology of uniparental disomy. *Hum. Genet.* **100:** 415–419.

Ehrlich M. 2003. The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clin. Immunol.* **109:** 17–28.

Ehrlich M., Buchanan K.L., Tsien F., Jiang G., Sun B., Uicker W., Weemaes C.M., Smeets D., Sperling K., Belohradsky B.H., et al. 2001. DNA methyltransferase 3B mutations linked to the ICF syndrome cause dysregulation of lymphogenesis genes. *Hum. Mol. Genet.* **10:** 2917–2931.

Engel E. 1980. A new genetic concept: Uniparental disomy and its potential effect, isodisomy. *Am. J. Med. Genet.* **6:** 137–143.

Fattal-Valevski A., Bassan H., Korman S.H., Lerman-Sagie T., Gutman A., and Harel S. 2000. Methylenetetrahydrofolate reductase deficiency: Importance of early diagnosis. *J. Child Neurol.* **15:** 539–543.

Feng Y., Absher D., Eberhart D.E., Brown V., Malter H.E., and Warren S.T. 1997. FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association. *Mol Cell* **1:** 109–118.

Flint J., Thomas K., Micklem G., Raynham H., Clark K., Doggett N.A., King A., and Higgs D.R. 1997. The relationship between chromosome structure and function at a human telomeric region. *Nat. Genet.* **15:** 252–257.

Forrester W.C., Epner E., Driscoll M.C., Enver T., Brice M., Papayannopoulou T., and Groudine M. 1990. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. *Genes Dev.* **4:** 1637–1649.

Fraga M.F., Ballestar E., Paz M.F., Ropero S., Setien F., Ballestar M.L., Heine-Suner D., Cigudosa J.C., Urioste M., Benitez J., et al. 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci.* **102:** 10604–10609.

Gabellini D., Green M.R., and Tupler R. 2002. Inappropriate gene activation in FSHD: A repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* **110:** 339–348.

Gibbons R.J., Picketts D.J., Villard L., and Higgs D.R. 1995. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell* **80:** 837–845.

Gibbons R.J., Pellagatti A., Garrick D., Wood W.G., Malik N., Ayyub H., Langford C., Boultwood J., Wainscoat J.S., and Higgs D.R. 2003. Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the α-thalassemia myelodysplasia syndrome (ATMDS). *Nat. Genet.* **34:** 446–449.

Gicquel C., Rossignol S., Cabrol S., Houang M., Steunou V., Barbu V., Danton F., Thibaud N., Le Merrer M., Burglen L., et al. 2005. Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. *Nat. Genet.* **37:** 1003–1007.

Gowher H. and Jeltsch A. 2002. Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J. Biol. Chem.* **277:** 20409–20414.

Goyette P., Sumner J.S., Milos R., Duncan A.M., Rosenblatt D.S., Matthews R.G., and Rozen R. 1994. Human methylenetetrahydrofolate reductase: Isolation of cDNA, mapping and mutation identification. *Nat. Genet.* **7:** 195–200.

Grosveld F. 1999. Activation by locus control regions? *Curr. Opin. Genet. Dev.* **9:** 152–157.

Hagberg B., Aicardi J., Dias K., and Ramos O. 1983. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: Report of 35 cases. *Ann. Neurol.* **14:** 471–479.

Hagerman P.J. and Hagerman R.J. 2004. The fragile-X premutation: A maturing perspective. *Am. J. Hum. Genet.* **74:** 805–816.

Hagerman R.J., Van Housen K., Smith A.C., and McGavran L. 1984. Consideration of connective tissue dysfunction in the fragile X syndrome. *Am. J. Med. Genet.* **17:** 111–121.

Handa V., Saha T., and Usdin K. 2003. The fragile X syndrome repeats form RNA hairpins that do not activate the interferon-inducible protein kinase, PKR, but are cut by Dicer. *Nucleic Acids Res.* **31:** 6243–6248.

Hansen R.S., Wijmenga C., Luo P., Stanek A.M., Canfield T.K., Weemaes C.M., and Gartler S.M. 1999. The *DNMT3B* DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc. Natl. Acad. Sci.* **96:** 14412–14417.

Harikrishnan K.N., Chow M.Z., Baker E.K., Pal S., Bassal S., Brasacchio D., Wang L., Craig J.M., Jones P.L., Sif S., and El-Osta A. 2005. Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat. Genet.* **37:** 254–264.

Hark A.T., Schoenherr C.J., Katz D.J., Ingram R.S., Levorse J.M., and Tilghman S.M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* **405:** 486–489.

Harrison C.J., Jack E.M., Allen T.D., and Harris R. 1983. The fragile X: A scanning electron microscope study. *J. Med. Genet.* **20:** 280–285.

Hasegawa T., Hara M., Ando M., Osawa M., Fukuyama Y., Takahashi M., and Yamada K. 1984. Cytogenetic studies of familial Prader-Willi syndrome. *Hum. Genet.* **65:** 325–330.

Hayward B.E., Kamiya M., Strain L., Moran V., Campbell R., Hayashizaki Y., and Bonthron D.T. 1998. The human *GNAS1* gene is imprinted and encodes distinct paternally and biallelically expressed G proteins. *Proc. Natl. Acad. Sci.* **95:** 10038–10043.

Henry I., Bonaiti-Pellie C., Chehensse V., Beldjord C., Schwartz C., Utermann G., and Junien C. 1991. Uniparental paternal disomy in a genetic cancer-predisposing syndrome. *Nature* **351:** 665–667.

Hoal-van Helden E.G. and van Helden P.D. 1989. Age-related methylation changes in DNA may reflect the proliferative potential of organs. *Mutat. Res.* **219:** 263–266.

Horike S., Cai S., Miyano M., Cheng J.F., and Kohwi-Shigematsu T. 2005. Loss of silent-chromatin looping and impaired imprinting of *DLX5* in Rett syndrome. *Nat. Genet.* **37:** 31–40.

Hubbard V.S., Davis P.B., di Sant'Agnese P.A., Gorden P., and Schwartz R.H. 1980. Isolated growth hormone deficiency and cystic fibrosis: A report of two cases. *Am. J. Dis. Child* **134:** 317–319.

Ingrosso D., Cimmino A., Perna A.F., Masella L., De Santo N.G., De Bonis M.L., Vacca M., D'Esposito M., D'Urso M., Galletti P., and Zappia V. 2003. Folate treatment and unbalanced methylation and changes of allelic expression induced by hyperhomocysteinaemia in patients with uraemia. *Lancet* **361:** 1693–1699.

Ishizuka A., Siomi M.C., and Siomi H. 2002. A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev.* **16:** 2497–2508.

Issa J.P., Ottaviano Y.L., Celano P., Hamilton S.R., Davidson N.E., and Baylin S.B. 1994. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. *Nat. Genet.* **7:** 536–540.

Jeanpierre M., Turleau C., Aurias A., Prieur M., Ledeist F., Fischer A., and Viegas-Pequignot E. 1993. An embryonic-like methylation

pattern of classical satellite DNA is observed in ICF syndrome. *Hum. Mol. Genet.* **2:** 731–735.

Jiang Y.H., Armstrong D., Albrecht U., Atkins C.M., Noebels J.L., Eichele G., Sweatt J.D., and Beaudet A.L. 1998. Mutation of the Angelman ubiquitin ligase in mice causes increased cytoplasmic p53 and deficits of contextual learning and long-term potentiation (see comments). *Neuron* **21:** 799–811.

Jin P., Alisch R.S., and Warren S.T. 2004a. RNA and microRNAs in fragile X mental retardation. *Nat. Cell Biol.* **6:** 1048–1053.

Jin P., Zarnescu D.C., Zhang F., Pearson C.E., Lucchesi J.C., Moses K., and Warren S.T. 2003. RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in *Drosophila*. *Neuron* **39:** 739–747.

Jin P., Zarnescu D.C., Ceman S., Nakamoto M., Mowrey J., Jongens T.A., Nelson D.L., Moses K., and Warren S.T. 2004b. Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. *Nat. Neurosci.* **7:** 113–117.

Joyce J.A., Lam W.K., Catchpoole D.J., Jenks P., Reik W., Maher E.R., and Schofield P.N. 1997. Imprinting of IGF2 and H19: Lack of reciprocity in sporadic Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* **6:** 1543–1548.

Kaufmann W.E., Jarrar M.H., Wang J.S., Lee Y.J., Reddy S., Bibat G., and Naidu S. 2005. Histone modifications in Rett syndrome lymphocytes: A preliminary evaluation. *Brain Dev.* **27:** 331–339.

Kioussis D., Vanin E., deLange T., Flavell R.A., and Grosveld F.G. 1983. Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. *Nature* **306:** 662–666.

Kishi N. and Macklis J.D. 2004. MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. *Mol. Cell. Neurosci.* **27:** 306–321.

Kishino T., Lalande M., and Wagstaff J. 1997. *UBE3A/E6-AP* mutations cause Angelman syndrome. *Nat. Genet.* **15:** 70–73.

Kondo T., Bobek M.P., Kuick R., Lamb B., Zhu X., Narayan A., Bourc'his D., Viegas-Pequignot E., Ehrlich M., and Hanash S.M. 2000. Whole-genome methylation scan in ICF syndrome: Hypomethylation of non-satellite DNA repeats *D4Z4* and *NBL2*. *Hum. Mol. Genet.* **9:** 597–604.

Korenke G.C., Fuchs S., Krasemann E., Doerr H.G., Wilichowski E., Hunneman D.H., and Hanefeld F. 1996. Cerebral adrenoleukodystrophy (ALD) in only one of monozygotic twins with an identical ALD genotype. *Ann. Neurol.* **40:** 254–257.

Kwok R.P., Lundblad J.R., Chrivia J.C., Richards J.P., Bachinger H.P., Brennan R.G., Roberts S.G., Green M.R., and Goodman R.H. 1994. Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature* **370:** 223–226.

Laggerbauer B., Ostareck D., Keidel E.M., Ostareck-Lederer A., and Fischer U. 2001. Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum. Mol. Genet.* **10:** 329–338.

Ledbetter D.H., Riccardi V.M., Airhart S.D., Strobel R.J., Keenan B.S., and Crawford J.D. 1981. Deletions of chromosome 15 as a cause of the Prader-Willi syndrome. *N. Engl. J. Med.* **304:** 325–329.

Lee M.P., DeBaun M.R., Mitsuya K., Galonek H.L., Brandenburg S., Oshimura M., and Feinberg A.P. 1999. Loss of imprinting of a paternally expressed transcript, with antisense orientation to $K_vLQT1$, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc. Natl. Acad. Sci.* **96:** 5203–5208.

Lemmers R.J., de Kievit P., Sandkuijl L., Padberg G.W., van Ommen G.J., Frants R.R., and van der Maarel S.M. 2002. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat. Genet.* **32:** 235–236.

Lewis J.D., Meehan R.R., Henzel W.J., Maurer-Fogy I., Jeppesen P., Klein F., and Bird A. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69:** 905–914.

Li Z., Zhang Y., Ku L., Wilkinson K.D., Warren S.T., and Feng Y. 2001. The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Res.* **29:** 2276–2283.

Liu D., Diorio J., Tannenbaum B., Caldji C., Francis D., Freedman A., Sharma S., Pearson D., Plotsky P.M., and Meaney M.J. 1997. Maternal care, hippocampal glucocorticoid receptors, and hypothalamic-pituitary-adrenal responses to stress. *Science* **277:** 1659–1662.

Lubs H.A. 1969. A marker X chromosome. *Am. J. Hum. Genet.* **21:** 231–244.

Ludwig M., Katalinic A., Gross S., Sutcliffe A., Varon R., and Horsthemke B. 2005. Increased prevalence of imprinting defects in patients with Angelman syndrome born to subfertile couples. *J. Med. Genet.* **42:** 289–291.

Ma J., Stampfer M.J., Giovannucci E., Artigas C., Hunter D.J., Fuchs C., Willett W.C., Selhub J., Hennekens C.H., and Rozen R. 1997. Methylenetetrahydrofolate reductase polymorphism, dietary interactions, and risk of colorectal cancer. *Cancer Res.* **57:** 1098–1102.

Magenis R.E., Brown M.G., Lacy D.A., Budden S., and LaFranchi S. 1987. Is Angelman syndrome an alternate result of del(15)(q11q13)? *Am. J. Med. Genet.* **28:** 829–838.

Martin J. and Bell J. 1943. A pedigree of mental defect showing sex-linkage. *Arch. Neurol. Psychiat.* **6:** 154–157.

Martinowich K., Hattori D., Wu H., Fouse S., He F., Hu Y., Fan G., and Sun Y.E. 2003. DNA methylation-related chromatin remodeling in activity-dependent *Bdnf* gene regulation. *Science* **302:** 890–893.

Mathews K.D. 2003. Muscular dystrophy overview: Genetics and diagnosis. *Neurol. Clin.* **21:** 795–816.

Matsuura T., Sutcliffe J.S., Fang P., Galjaard R.J., Jiang Y.H., Benton C.S., Rommens J.M., and Beaudet A.L. 1997. De novo truncating mutations in E6-AP ubiquitin-protein ligase gene (*UBE3A*) in Angelman syndrome. *Nat. Genet.* **15:** 74–77.

Mayr B. and Montminy M. 2001. Transcriptional regulation by the phosphorylation-dependent factor CREB. *Nat. Rev. Mol. Cell. Biol.* **2:** 599–609.

Mays-Hoopes L.L. 1989. Age-related changes in DNA methylation: Do they represent continued developmental changes? *Int. Rev. Cytol.* **114:** 181–220.

McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M., Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J., Buckle V.J., et al. 1999. Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc. Natl. Acad. Sci.* **96:** 13983–13988.

Meguro M., Mitsuya K., Nomura N., Kohda M., Kashiwagi A., Nishigaki R., Yoshioka H., Nakao M., Oishi M., and Oshimura M. 2001. Large-scale evaluation of imprinting status in the Prader-Willi syndrome region: An imprinted direct repeat cluster resembling small nucleolar RNA genes. *Hum. Mol. Genet.* **10:** 383–394.

Meins M., Lehmann J., Gerresheim F., Herchenbach J., Hagedorn M., Hameister K., and Epplen J.T. 2005. Submicroscopic duplication in Xq28 causes increased expression of the *MECP2* gene in a boy with severe mental retardation and features of Rett syndrome. *J. Med. Genet.* **42:** e12.

Meloni I., Bruttini M., Longo I., Mari F., Rizzolio F., D'Adamo P., Denvriendt K., Fryns J.P., Toniolo D., and Renieri A. 2000. A mutation in the Rett syndrome gene, *MECP2*, causes X-linked mental retardation and progressive spasticity in males. *Am. J. Hum. Genet.* **67:** 982–985.

Mullaney B.C., Johnston M.V., and Blue M.E. 2004. Developmental expression of methyl-CpG binding protein 2 is dynamically regulated in the rodent brain. *Neuroscience* **123:** 939–949.

Murata T., Kurokawa R., Krones A., Tatsumi K., Ishii M., Taki T., Masuno M., Ohashi H., Yanagisawa M., Rosenfeld M.G., et al. 2001. Defect of histone acetyltransferase activity of the nuclear transcriptional coactivator CBP in Rubinstein-Taybi syndrome. *Hum. Mol. Genet.* **10:** 1071–1076.

Nan X., Campoy F.J., and Bird A. 1997. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88:** 471–481.

Neul J.L. and Zoghbi H.Y. 2004. Rett syndrome: A prototypical neurodevelopmental disorder. *Neuroscientist* **10:** 118–128.

Nicholls R.D., Knoll J.H.M., Butler M.G., Karam S., and Lalande M. 1989. Genetic imprinting suggested by maternal heterodisomy in non-deletion Prader-Willi syndrome. *Nature* **342:** 281–285.

Nuber U.A., Kriaucionis S., Roloff T.C., Guy J., Selfridge J., Steinhoff C., Schulz R., Lipkowitz B., Ropers H.H., Holmes M.C., and Bird A. 2005. Up-regulation of glucocorticoid-regulated genes in a mouse

model of Rett syndrome. *Hum. Mol. Genet.* **14:** 2247–2256.

Ogryzko V.V., Schiltz R.L., Russanova V., Howard B.H., and Nakatani Y. 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87:** 953–959.

Ohta T., Gray T.A., Rogan P.K., Buiting K., Gabriel J.M., Saitoh S., Muralidhar B., Bilienska B., Krajewska-Walasek M., Driscoll D.J., et al. 1999. Imprinting-mutation mechanisms in Prader-Willi syndrome. *Am. J. Hum. Genet.* **64:** 397–413.

Okano M., Bell D.W., Haber D.A., and Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99:** 247–257.

Orstavik K.H., Eiklid K., van der Hagen C.B., Spetalen S., Kierulf K., Skjeldal O., and Buiting K. 2003. Another case of imprinting defect in a girl with Angelman syndrome who was conceived by intracytoplasmic semen injection. *Am. J. Hum. Genet.* **72:** 218–219.

Petrij F., Giles R.H., Dauwerse H.G., Saris J.J., Hennekam R.C., Masuno M., Tommerup N., van Ommen G.J., Goodman R.H., Peters D.J., et al. 1995. Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **376:** 348–351.

Petronis A. 2004. The origin of schizophrenia: Genetic thesis, epigenetic antithesis, and resolving synthesis. *Biol. Psychiatry* **55:** 965–970.

Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W., and Gibbons R.J. 1996. ATRX encodes a novel member of the SNF2 family of proteins: Mutations point to a common mechanism underlying the ATR-X syndrome. *Hum. Mol. Genet.* **5:** 1899–1907.

Pieretti M., Zhang F., Fu Y.-H., Warren S.T., Oostra B.A., Caskey C.T., and Nelson D.L. 1991. Absence of expression of the *FMR-1* gene in fragile X syndrome. *Cell* **66:** 817–822.

Ping A.J., Reeve A.E., Law D.J., Young M.R., Boehnke M., and Feinberg A.P. 1989. Genetic linkage of Beckwith-Wiedemann syndrome to 11p15. *Am. J. Hum. Genet.* **44:** 720–773.

Prawitt D., Enklaar T., Gartner-Rupprecht B., Spangenberg C., Oswald M., Lausch E., Schmidtke P., Reutzel D., Fees S., Lucito R., et al. 2005. Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms' tumor. *Proc. Natl. Acad. Sci.* **102:** 4085–4090.

Rampersaud G.C., Kauwell G.P., Hutson A.D., Cerda J.J., and Bailey L.B. 2000. Genomic DNA methylation decreases in response to moderate folate depletion in elderly women. *Am. J. Clin. Nutr.* **72:** 998–1003.

Reik W. 1989. Genomic imprinting and genetic disorders in man. *Trends Genet.* **5:** 331–336.

Reynolds E.H., Carney M.W., and Toone B.K. 1984. Methylation and mood. *Lancet* **2:** 196–198.

Richards B.W., Sylvester P.E., and Brooker C. 1981. Fragile X-linked mental retardation: The Martin-Bell syndrome. *J. Ment. Defic. Res.* **25:** 253–256.

Roelfsema J.H., White S.J., Ariyurek Y., Bartholdi D., Niedrist D., Papadia F., Bacino C.A., den Dunnen J.T., van Ommen G.J., Breuning M.H., et al. 2005. Genetic heterogeneity in Rubinstein-Taybi syndrome: Mutations in both the *CBP* and *EP300* genes cause disease. *Am. J. Hum. Genet.* **76:** 572–580.

Rougeulle C., Cardoso C., Fontes M., Colleaux L., and Lalande M. 1998. An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat. Genet.* **19:** 15–16.

Rozen R. 1996. Molecular genetics of methylenetetrahydrofolate reductase deficiency. *J. Inherit. Metab. Dis.* **19:** 589–594.

Runte M., Varon R., Horn D., Horsthemke B., and Buiting K. 2005. Exclusion of the C/D box snoRNA gene cluster *HBII-52* from a major role in Prader-Willi syndrome. *Hum. Genet.* **116:** 228–230.

Schimke R.N., Horton W.A., and King C.R. 1971. Chondroitin-6-sulphaturia, defective cellular immunity, and nephrotic syndrome. *Lancet* **2:** 1088–1089.

Schule B., Albalwi M., Northrop E., Francis D.I., Rowell M., Slater H.R., Gardner R.J., and Francke U. 2005. Molecular breakpoint cloning and gene expression studies of a novel translocation t(4;15)(q27;q11.2) associated with Prader-Willi syndrome. *BMC Med. Genet.* **6:** 18.

Schwahn B. and Rozen R. 2001. Polymorphisms in the methylenetetrahydrofolate reductase gene: Clinical consequences. *Am. J. Pharmacogenomics* **1:** 189–201.

Shahbazian M.D., Antalffy B., Armstrong D.L., and Zoghbi H.Y. 2002a. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum. Mol. Genet.* **11:** 115–124.

Shahbazian M., Young J., Yuva-Paylor L., Spencer C., Antalffy B., Noebels J., Armstrong D., Paylor R., and Zoghbi H. 2002b. Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron* **35:** 243–254.

Smeets D.F., Moog U., Weemaes C.M., Vaes-Peeters G., Merkx G.F., Niehof J.P., and Hamers G. 1994. ICF syndrome: A new case and review of the literature. *Hum. Genet.* **94:** 240–246.

Smeets D.F., Hamel B.C., Nelen M.R., Smeets H.J., Bollen J.H., Smits A.P., Ropers H.H., and van Oost B.A. 1992. Prader-Willi syndrome and Angelman syndrome in cousins from a family with a translocation between chromosomes 6 and 15. *N. Engl. J. Med.* **326:** 807–811.

Smilinich N.J., Day C.D., Fitzpatrick G.V., Caldwell G.M., Lossie A.C., Cooper P.R., Smallwood A.C., Joyce J.A., Schofield P.N., Reik W., et al. 1999. A maternally methylated CpG island in *KvLQT1* is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proc. Natl. Acad. Sci.* **96:** 8064–8069.

Spence J.E., Perciaccante R.G., Greig G.M., Willard H.F., Ledbetter D.H., Hejtmancik J.F., Pollack M.S., O'Brien W.E., and Beaudet A.L. 1988. Uniparental disomy as a mechanism for human genetic disease. *Am. J. Hum. Genet.* **42:** 217–226.

Spranger J., Hinkel G.K., Stoss H., Thoenes W., Wargowski D., and Zepp F. 1991. Schimke immuno-osseous dysplasia: A newly recognized multisystem disease. *J. Pediatr.* **119:** 64–72.

Sun F.L., Dean W.L., Kelsey G., Allen N.D., and Reik W. 1997. Transactivation of *Igf2* in a mouse model of Beckwith-Wiedemann syndrome. *Nature* **389:** 809–815.

Suter C.M., Martin D.I., and Ward R.L. 2004. Germline epimutation of MLH1 in individuals with multiple cancers. *Nat. Genet.* **36:** 497–501.

Sutherland G.R. 1977. Fragile sites on human chromosomes: Demonstration of their dependence on the type of tissue culture medium. *Science* **197:** 265–266.

Taha D., Boerkoel C.F., Balfe J.W., Khalifah M., Sloan E.A., Barbar M., Haider A., and Kanaan H. 2004. Fatal lymphoproliferative disorder in a child with Schimke immuno-osseous dysplasia. *Am. J. Med. Genet. A* **131:** 194–199.

Tommerup N., van der Hagen C.B., and Heiberg A. 1992. Tentative assignment of a locus for Rubinstein-Taybi syndrome to 16p13.3 by a de novo reciprocal translocation, t(7;16)(q34;p13.3). *Am. J. Med. Genet.* **44:** 237–241.

Tsai T.F., Jiang Y.H., Bressler J., Armstrong D., and Beaudet A.L. 1999. Paternal deletion from Snrpn to Ube3a in the mouse causes hypotonia, growth retardation and partial lethality and provides evidence for a gene contributing to Prader-Willi syndrome. *Hum. Mol. Genet.* **8:** 1357–1364.

Tuck-Muller C.M., Narayan A., Tsien F., Smeets D.F., Sawyer J., Fiala E.S., Sohn O.S., and Ehrlich M. 2000. DNA hypomethylation and unusual chromosome instability in cell lines from ICF syndrome patients. *Cytogenet. Cell Genet.* **89:** 121–128.

Tupler R. and Gabellini D. 2004. Molecular basis of facioscapulohumeral muscular dystrophy. *Cell. Mol. Life Sci.* **61:** 557–566.

van Deutekom J.C., Wijmenga C., van Tienhoven E.A., Gruter A.M., Hewitt J.E., Padberg G.W., van Ommen G.J., Hofker M.H., and Frants R.R. 1993. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.* **2:** 2037–2042.

Van Esch H., Bauters M., Ignatius J., Jansen M., Raynaud M., Hollanders K., Lugtenberg D., Bienvenu T., Jensen L.R., Gecz J., et al. 2005. Duplication of the *MECP2* region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am. J. Hum. Genet.* **77:** 442–453.

Verkerk A.J.M.H., Pieretti M., Sutcliffe J.S., Fu Y.-H., Kuhl D.P.A., Pizutti A., Reiner O., Richards S., Victoria M.F., Zhang R., et al. 1991. Identification of a gene (FMR-1) containing a CGG repeat coincident

with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65:** 905–914.

Villard L., Gecz J., Mattei J.F., Fontes M., Saugier-Veber P., Munnich A., and Lyonnet S. 1996. XNP mutation in a large family with Juberg-Marsidi syndrome. *Nat. Genet.* **12:** 359–360.

Wan M., Zhao K., Lee S.S., and Francke U. 2001. *MECP2* truncating mutations cause histone H4 hyperacetylation in Rett syndrome. *Hum. Mol. Genet.* **10:** 1085–1092.

Wan M., Lee S.S., Zhang X., Houwink-Manville I., Song H.R., Amir R.E., Budden S., Naidu S., Pereira J.L., Lo I.F., et al. 1999. Rett syndrome and beyond: Recurrent spontaneous and familial *MECP2* mutations at CpG hotspots. *Am. J. Hum. Genet.* **65:** 1520–1529.

Warren S.T. and Sherman S.L. 2001. The fragile X syndrome. In *The metabolic and molecular bases of inherited disease*, 8th edition (ed. C.R. Scriver et al.), vol. 1, pp. 1257–1289. McGraw-Hill, New York.

Waterland R.A. and Jirtle R.L. 2003. Transposable elements: Targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* **23:** 5293–5300.

Weatherall D.J., Clegg M.B., Higgs D.R., and Wood W.G. 2001. The hemoglobinopathies. In *The metabolic & molecular bases of inherited disease*, 8th edition (ed. C.R. Scriver et al.), pp. 4571–4636. McGraw-Hill, New York.

Weaver I.C., Cervoni N., Champagne F.A., D'Alessio A.C., Sharma S., Seckl J.R., Dymov S., Szyf M., and Meaney M.J. 2004. Epigenetic programming by maternal behavior. *Nat. Neurosci.* **7:** 847–854.

Weiler I.J. and Greenough W.T. 1999. Synaptic synthesis of the Fragile X protein: Possible involvement in synapse maturation and elimination. *Am. J. Med. Genet.* **83:** 248–252.

Weksberg R., Smith A.C., Squire J., and Sadowski P. 2003. Beckwith-Wiedemann syndrome demonstrates a role for epigenetic control of normal development. *Hum. Mol. Genet.* **12:** R61–R68.

Weksberg R., Nishikawa J., Caluseriu O., Fei Y.L., Shuman C., Wei C., Steele L., Cameron J., Smith A., Ambus I., et al. 2001. Tumor development in the Beckwith-Wiedemann syndrome is associated with a variety of constitutional molecular 11p15 alterations including imprinting defects of *KCNQ1OT1*. *Hum. Mol. Genet.* **10:** 2989–3000.

Weksberg R., Teshima I., Williams B.R., Greenberg C.R., Pueschel S.M., Chernos J.E., Fowlow S.B., Hoyme E., Anderson I.J., Whiteman D.A., et al. 1993. Molecular characterization of cytogenetic alterations associated with the Beckwith-Wiedemann syndrome (BWS) phenotype refines the localization and suggests the gene for BWS is imprinted. *Hum. Mol. Genet.* **2:** 549–556.

Wijmenga C., Hewitt J.E., Sandkuijl L.A., Clark L.N., Wright T.J., Dauwerse H.G., Gruter A.M., Hofker M.H., Moerer P., Williamson R., et al. 1992. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.* **2:** 26–30.

Willemsen R., Hoogeveen-Westerveld M., Reis S., Holstege J., Severijnen L.A., Nieuwenhuizen I.M., Schrier M., Van Unen L., Tassone F., Hoogeveen A.T., et al. 2003. The *FMR1* CGG repeat mouse displays ubiquitin-positive intranuclear neuronal inclusions; implications for the cerebellar tremor/ataxia syndrome. *Hum. Mol. Genet.* **12:** 949–959.

Wolff G.L., Kodell R.L., Moore S.R., and Cooney C.A. 1998. Maternal epigenetics and methyl supplements affect *agouti* gene expression in $A^{vy}/a$ mice. *Faseb J.* **12:** 949–957.

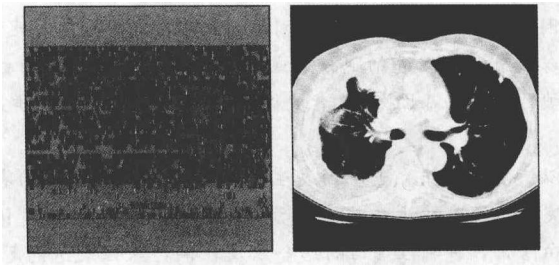
Xu G.L., Bestor T.H., Bourc'his D., Hsieh C.L., Tommerup N., Bugge M., Hulten M., Qu X., Russo J.J., and Viegas-Pequignot E. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402:** 187–191.

Yntema H.G., Poppelaars F.A., Derksen E., Oudakker A.R., van Roosmalen T., Jacobs A., Obbema H., Brunner H.G., Hamel B.C., and van Bokhoven H. 2002. Expanding phenotype of XNP mutations: Mild to moderate mental retardation. *Am. J. Med. Genet.* **110:** 243–247.

Young J.I., Hong E.P., Castle J., Crespo-Barreto J., Bowman A.B., Rose M.F., Kang D., Richman R., Johnson J.M., Berget S., and Zoghbi H.Y. 2005. Inaugural article: Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc. Natl. Acad. Sci.* **102:** 17551–17558.

Zeev B.B., Yaron Y., Schanen N.C., Wolf H., Brandt N., Ginot N., Shomrat R., and Orr-Urtreger A. 2002. Rett syndrome: Clinical manifestations in males with *MECP2* mutations. *J. Child Neurol.* **17:** 20–24.

# 第 24 章



# 癌症的表观遗传决定因素

Stephen B. Baylin[1] and Peter A. Jones[2]

[1]*Cancer Biology Program, The Sidney Kimmel Cancer Center, Johns Hopkins Medical Institutions, Baltimore, Maryland 21231*

[2]*Department of Urology, Biochemistry and Molecular Biology, USC/Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, California 90089-9181*

## 本章目录

1. 癌症的生物学基础
2. 染色质对于癌症的重要意义
3. DNA 甲基化在癌症中的作用
4. 癌症中基因启动子的高甲基化
   - 4.1 涉及的基因
   - 4.2 在癌症细胞中寻找新的表观沉默基因
   - 4.3 确认癌症中高甲基化基因的重要作用
5. 表观遗传基因沉默及其在癌症形成中的作用
6. 分子水平剖析表观遗传沉默的癌症基因
7. 概述癌症中表观遗传基因沉默的主要研究方向
8. 通过 DNA 甲基化检测癌症
9. 表观遗传治疗

参考文献

## 概要

癌症由可继承的基因调控错误引起，这些基因控制细胞在何时分裂、死亡以及迁移。在癌症发生过程中，有些基因可以被激活以促进细胞分裂或阻止细胞死亡，有些则被沉默从而不再能阻止以上进程的发生。前者被称为“癌基因”，后者被称为“抑癌基因”。这两类基因相互影响的结果最终导致癌症的发生。

有至少三种方式可以使基因沉默。①基因发生突变从而失去功能。②基因被完全缺失无法进行正常工作。③既没有突变也没有缺失的基因由于表观遗传的改变导致遗传的沉默。这种表观遗传的沉默可以涉及组蛋白修饰，抑制性蛋白的结合以及在调控基因表达的 CpG 岛区域上胞嘧啶的异常甲基化。

本章重点阐述第三种途径。正如本书所描述的，维持基因沉默的分子机制已经了解得非常透彻了。不仅如此，我们还知道表观遗传沉默对于癌症的预防、检测和治疗都有着深远的影响。我们已经有一些经过美国 FDA 批准的、可以在癌细胞中逆转表观遗传变化并且恢复基因表达的药物。由于 DNA 甲基化的改变可以被灵敏地检测到，许多检测早期癌症病变的策略依赖于对 DNA 甲基化变化的寻找。在人类癌症研究、检测、预防以及治疗中，表观遗传都有着非常重要的应用前景。

## 1. 癌症的生物学基础

癌症从根本上来说是基因表达的紊乱，使多细胞生物中受到复杂网络调节的基因被扰乱，导致细胞生长不受控制，而不顾及生物个体的总体需求。在癌症中，相关细胞调控途径被扰乱的概况已经取得了重大进展（表 24-1）。几乎在所有癌症中，都是因为这些有限的细胞调控通路受到了影响并被可继承地失活，这是推进了本领域发展的核心概念（Hanahan and Weinberg 2000）。直到最近几年，焦点仍集中在癌症的遗传学基础，特别是因突变导致的癌基因激活或抑癌基因的沉默。然而，从 20 世纪 90 年代中期开始，越来越多的数据表明由表观遗传改变调控的可继承的变化，可能也对人类所有癌症的演变起至关重要的作用（图 24-1）（Jones and Laird 1999；Jones and Baylin 2002；Herman and Baylin 2003）。这些数据，尤其是在癌症中显著改变的 DNA 和染色质甲基化模式，为癌症的研究、检测、治疗及预防提供了新的机遇。

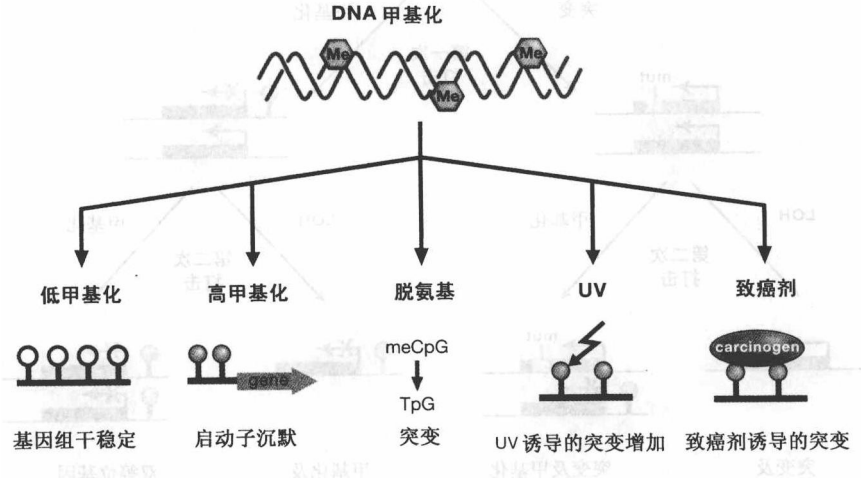



图 24-1　DNA 甲基化有关的表观遗传改变可以通过多种机制引起癌症

DNA 胞嘧啶甲基化丢失（低甲基化）导致基因组不稳定。基因启动子局部高甲基化可以引起可继承的沉默并使抑癌基因失活。此外，甲基化 CpG 位点是自发脱氨反应引起 C→T 的突变的热点。甲基化 CpG 位点也可以增加某些化学致癌剂与 DNA 的结合同时增加紫外诱导的突变几率。

**表 24-1　在人类癌症中被遗传或表观遗传机制破坏的重要细胞通路**

| 通路 | 遗传改变的例子 | 表观遗传改变的例子 |
|---|---|---|
| 生长自足信号 | *RAS* 基因突变 | *RASSF1A* 基因甲基化 |
| 抗生长信号的耐受 | TGFβ 受体基因突变 | TGFβ 受体下调 |
| 组织侵染和肿瘤转移 | *E-Cadherin* 基因突变 | *E-Cadherin* 基因启动子甲基化 |
| 无限复制潜能 | *p16* 和 *Rb* 基因突变 | 通过启动子甲基化沉默 *p16* 和 *Rb* 基因持续的血管生成 *thrombospondin-1* 沉默 |
| 规避凋亡 | *p53* 基因突变 | *DAP* 激酶，*ASC/TMS1* 以及 *HIC1* 甲基化 |
| DNA 修复能力 | *MLH1* 和 *MSH2* 基因突变 | *GST Pi*，*O6-MGMT*，*MLH1* 甲基化 |
| 基因组稳定性的监测 | *Chfr* 基因突变 | *Chfr* 基因甲基化 |
| 蛋白质泛素化功能 | *Chfr* 基因突变 | *Chfr* 基因甲基化 |

遗传和表观遗传异常可以通过两种不同的机制引起自我平衡通路的可继承性破坏。要么使癌基因被激活，要么使抑癌基因失活（Jones and Laird 1999；Hanahan and Weinberg 2000；Jones and Baylin 2002；Herman and Baylin 2003）。例如，单个基因（癌基因）如*RAS*基因的突变就常在人类癌症中被发现，这种突变提高了促进生长的基因产物的活性。这些突变通常是显性的并且可导致癌症发生。另一方面，抑癌基因的遗传突变和表观遗传沉默则通常是隐性的，需要两个等位基因同时被破坏才能产生新的表型。Knudson（2001）提出在恶性肿瘤细胞系中抑癌基因的两个等位基因都需要被破坏，这一观点得到了广泛的认可。人们认识到有三类“打击”可以通过不同的组合形式引起抑癌基因活性的完全消失。编码区域的突变，部分或整个基因的缺失以及表观遗传沉默可以互相作用导致重要的调控基因失活（图 24-2）。

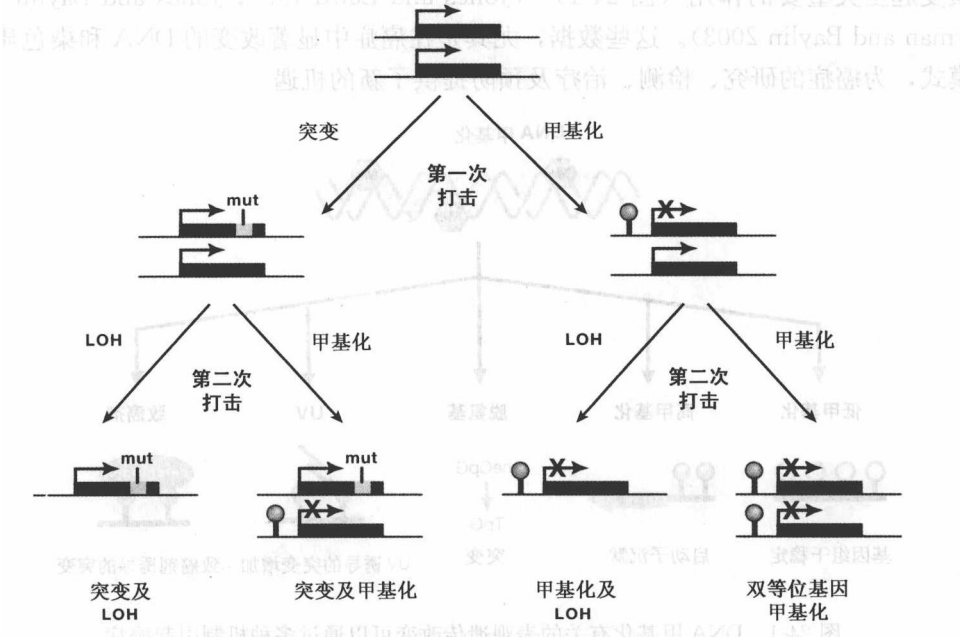



图 24-2 DNA 甲基化如何导致抑癌基因失活

两个深色的长方形代表抑癌基因的两个活跃等位基因。基因失活的第一步由局部突变所造成（左）或由 DNA 甲基化导致基因失活（右）。图示的第二次打击则来自于杂合性丢失（LOH）或累加的表观遗传事件引起的基因失活。这样，DNA 甲基化是满足 Knudson 假说的途径之一。

## 2. 染色质对于癌症的重要意义

尽管导致癌症的细胞调控通路中的核心分子的损伤研究已经取得了重大进展，但是病理学家对样品的核结构进行镜检仍然是癌症诊断的金标准。肉眼可以在单个细胞中准确的分辨核结构（或者主要来说是染色质状态）的变化，这是癌症最终的诊断标准。病理学家最初使用的标准是核的大小、核的形状、核膜是否皱缩，核仁是否显著，密集的

着色过深的染色质以及高核质比。在显微镜下清晰可见的这些结构特征（图 24-3）可能与染色质功能的深远改变有关，并且导致了基因表达状态以及染色体稳定性的改变。将这些在显微镜下可观察到的改变与本书所讨论的分子标记之间联系起来仍是癌症研究领域中的一个重大挑战。本章我们将回顾表观遗传标记，如 DNACpG 岛中胞嘧啶甲基化以及组蛋白修饰的改变，这些标记在癌症细胞中呈异常分布。这些标记越来越多地被联系到影响基因稳定和功能的可继承事件中，因此对癌症表型的产生起着重要作用。

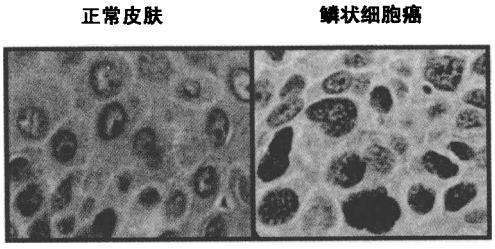



图 24-3　癌症细胞中染色质结构的改变

这是鳞状细胞癌患者皮肤的两张显微照片。图中对比了与肿瘤（右图）与相邻一厘米内的正常表皮细胞（左图），两图的放大倍数相同。染色质由于与苏木精的高亲和力而显示为紫色，其在癌症细胞中比在普通表皮细胞中显得更为粗糙且颗粒化。这种染色质染色特征的改变被病理学家用来作为诊断癌症的标准。

许多染色质修饰活性在人类癌症中的作用是已知的（Wolffe 2001）。例如，AML 综合征和 PML 综合征都是由染色体易位导致错误使用组蛋白去乙酰化酶（HDAC）导致的。在 PML 患者中，*PML* 基因与视黄酸受体基因（*RAR*）融合。该受体募集组蛋白去乙酰化酶和 DNA 甲基化酶，从而导致转录失活，这已被启动子实验的结果证实。这些数据暗示靶向性的染色质改变可能会导致抑癌基因的沉默，参与阻止细胞分化（Di Croce et al. 2002）。在 AML 患者中，转录因子 AML-1 的 DNA 结合结构域与一个名为 ETO 的蛋白融合在一起，ETO 与 HDAC 相互作用。错误的组蛋白去乙酰化抑制细胞分化并导致错误的基因抑制最终导致白血病（Amann et al. 2001）。目前只有两个直接证据证明染色质修饰参与癌基因表型的形成。然而越来越清楚的是染色质修饰可以直接或间接地改变胞嘧啶甲基化的模式，这种 DNA 表观遗传的改变可以起始或保持重要基因处于沉默状态，从而导致在重要细胞通路中可继承的变化。

## 3. DNA 甲基化在癌症中的作用

除了 4 种直接进入 DNA 的碱基外，在 DNA 中还含有 5-甲基胞嘧啶，这一发现很快导致人们提出了 DNA 甲基化的改变可能致癌的假说（表 24-2）。在过去的 40 多年里，许多研究都表明在癌细胞和正常细胞中 DNA 5-甲基胞嘧啶的分布模式发生了改变。在这些细胞中，CpG 甲基化至少通过三种主要途径导致癌的表型。这包括癌基因的低甲基化、抑癌基因启动子的高甲基化以及直接的诱变（图 24-1）（Jones and Laird 1999；Jones and Baylin 2002；Herman and Baylin 2003）。尽管这些变化中的任何一个都

能单独引起人的癌变，但这三种途径更可能是同时发生，这表明改变自我平衡通路的表观遗传机制对癌症发生起着至关重要的作用。

**表 24-2 阐明癌症中 DNA 甲基化功能的大事记**

| 现象 | 参考文献 |
| --- | --- |
| “甲基化酶致癌性”假说 | Srinivasan and Borek(1964) |
| 动物肿瘤中 5-甲基胞嘧啶水平下降 | Lapeyre and Becker(1979) |
| 5-氮杂胞苷和 5-氮杂脱氧胞苷抑制甲基化并激活基因 | Jones and Taylor(1980) |
| 人类肿瘤中基因组水平和特殊基因上甲基化程度的降低 | Ehrlich et al. (1982); Feinberg and Vogelstein (1983); Flatau et al. (1983) |
| DNA 甲基化抑制剂可以改变恶性表型 | Frost et al. (1984) |
| 癌症中 CpG 岛的甲基化 | Baylin et al. (1987) |
| *P53* 的突变热点是甲基化 CpG 位点 | Rideout et al. (1990) |
| 视网膜母细胞瘤抑制基因(*RB*)的等位基因专一性甲基化 | Sakai et al. (1991) |
| 癌症中基因印记的缺失 | Rainier et al. (1993) |
| 随年龄增加的 CpG 岛高甲基化 | Issa et al. (1994) |
| 低甲基化水平的小鼠较少发生肿瘤 | Laird et al. (1995) |
| DNA 甲基化抑制剂和组蛋白去乙酰化酶抑制剂联合使用导致快速鉴定抑癌基因 | Suzuki et al. (2002); Yamashita et al. (2002) |
| 体细胞内 DNA 修复基因(MLH1)被甲基化 | Gazzoli et al. (2002) |
| 5-氮杂胞苷被 FDA 许可用于脊髓发育不良综合征的治疗 | Kaminskas et al. (2005) |

在癌症细胞中，最先知道也最明显的 DNA 甲基化模式变化是该修饰在整体水平上的下降，这会导致基因组的不稳定性（见第 18 章）。这是人类癌症的一个非常著名的标志（Feinberg and Vogelstein 1983; Feinberg et al. 1988; Jones and Laird 1999; Jones and Baylin 2002; Herman and Baylin 2003）。最近，大量的数据表明癌症相关基因的 5′区 CpG 岛甲基化的异常是转录沉默不可缺少的，这为抑癌基因功能失活提供了另外一种机制（Jones and Laird 1999; Jones and Baylin 2002; Herman and Baylin 2003）。最终，除了前面提到的胞嘧啶甲基化在基因组不稳定性和基因沉默中的作用外，5-甲基胞嘧啶自身的高度不稳定性使其成为一种可诱变因子。MeCpG 通过置换突变成为 TpG 可以直接导致癌症（Rideout et al. 1990）。事实上这些修饰在癌症中具有普遍性，且已知具有直接的致癌作用，因此表观遗传变化也许能够作为治疗的靶点（Egger et al. 2004）。

DNA 胞嘧啶甲基化现在被认为在人类致癌过程中起到了重要作用。几乎所有的人类基因在其编码区都含有甲基化的胞嘧啶，已知有时这些区域对于形成可引起疾病的突变起到了促进作用。胞嘧啶环上第五位 C 甲基化可以提高双链 DNA 碱基脱氨基的频率。然而，5-甲基胞嘧啶的脱氨基产物是胸腺嘧啶而不是尿嘧啶（见图 3-12）。随后的 DNA 修复机制对于 DNA 中脱氨基引起的错配修复效率较低。在人类生殖系细胞中，甲基化 CpG 位点导致了所有置换突变的 1/3 以上（Rideout et al. 1990）。对引起癌症的

基因也是如此，如 *p53*（Rideout et al. 1990）。更让人吃惊的观察是发现这种机制同样会在体细胞组织内导致大量的突变从而导致抑癌基因失活。例如，在散发性结肠直肠癌中所需要的 *p53* 突变中，有 50%发生在甲基化胞嘧啶位点（Greenblatt et al. 1994）。因此，DNA 甲基化酶产生的甲基化由于这种内源机制会大大提高患癌的风险。

甲基化胞嘧啶被证实更易在 DNA 和致癌物（如香烟中的苯并芘）间形成致癌性加成物。本例中，甲基化胞嘧啶会更多地与相邻鸟嘌呤和苯并芘二醇环氧化合物形成致癌加成物，从而导致吸烟者肺部 CpG 位点的突变增加（Greenblatt et al. 1994；Pfeifer et al. 2000）。有趣的是，在阳光暴晒下的皮肤中甲基化也可以改变 *p53* 基因的突变率（Greenblat et al. 1994；Pfeifer et al. 2000）。这是因为甲基基团将胞嘧啶的吸光光谱改变，进入到了阳光的光谱范围，所以在皮肤细胞中会形成大量的嘧啶二聚体。因此，DNA 表观遗传修饰不仅可以增加自发突变的形成，同时也可以影响 DNA 与致癌物质和紫外线的作用方式（Pfeifer et al. 2000）。

在动物和人类肿瘤中很早就发现了 DNA 低甲基化现象（表 24-2），DNA 低甲基化会影响染色体的稳定性并且增加非整倍性。染色体组的不稳定性是癌症的标志性标记，DNA 卫星序列以及其他序列低甲基化导致的染色体脆性会通过增加基因组不稳定性导致癌症发生（Narayan et al. 1998；Gaudet et al. 2003）。DNA 低甲基化导致不稳定性的确切机制还不清楚，但很可能是改变了 DNA 和蛋白质的相互作用。

## 4. 癌症中基因启动子的高甲基化

### 4.1　涉及的基因

抑癌基因启动子高甲基化导致的癌症是研究最为透彻的 DNA 甲基化导致癌症的机理。确切的高甲基化导致癌症的机制以下将详细讨论。然而，非常清楚的是这个重要的途径会导致抑癌基因可继承地沉默（Jones and Laird 1999；Jones and Baylin 2002；Herman and Baylin 2003）。通常，DNA 高甲基化发生在富含 CpG 的区域或 CpG 岛，这些区域位于癌症细胞被异常沉默的基因的转录起始位点或其附近。这对于我们认识基因起始位点周边的 CpG 岛胞嘧啶甲基化的作用是至关重要的，因为发生在基因内的相同的 DNA 修饰与转录状态间没有相关性（Jones 1999）。

受以上原因被破坏转录的癌症相关基因一直在稳步增加。正如以往回顾的，涉及的这些基因存在于所有的染色体位点（Jones and Baylin 2002）。事实上，在人类肿瘤中，发生表观遗传改变的基因要比常发生突变的基因更多。如前所述（表 24-1），通过 CpG 甲基化使抑癌基因功能消失，事实上影响了所有的已知通路。为了解这些基因对于肿瘤发生进程的重要性以及在该领域未来面临的挑战，这些基因可被分成三类。

第一类基因包括被用来定义启动子高甲基化以及基因沉默作为在癌症中使抑癌基因功能缺失的重要机理的基因（表 24-3）。这些是被承认的经典的抑癌基因，当它们在家系的生殖系细胞中发生突变时，会引起可继承的癌症发生（Jones and Laird 1999；Jones and Baylin 2002；Herman and Baylin 2003）。它们在散发性的癌症中也会发生突

变，且经常在一个或两个等位基因上发生高甲基化（Jones and Laird 1999；Jones and Baylin 2002；Herman and Baylin 2003）。此外，这些基因启动子的高甲基化在 Knudson 的假说中起到“第二次打击”的作用，在家系性的肿瘤中这通常与基因的第二拷贝功能性失活相关，第一次打击是生殖细胞内的突变（Grady et al. 2000；Esteller et al. 2001b）。在某些例子中，培养的肿瘤细胞经 5-氮杂胞苷诱导可重新激活这些基因，并恢复一些重要的抑癌基因在肿瘤发生过程中丢失的功能。在 *MLH1* 基因沉默的直肠癌细胞中错配修复功能的恢复是一个这样的例子（Herman et al. 1998）。

**表 24-3 各类高甲基化基因的发现**

1. 传统的抑癌基因在生殖细胞系中的突变可引起遗传性癌症（如 VHL，E-钙黏蛋白，p16Ink4a，MLH1，APC，Stk4，Rb）
2. 候选的抑癌基因（如 FHIT，RASSF1A，06-MGMT，Gst-Pi，GATA4 和 5，DAP 激酶）
3. 通过高甲基化随机筛选出的基因（HIC-1，SFRP1，2，4，5，BMP-3，SLC5A8，SSI1）

表观遗传沉默的第二类基因经体内功能实验被鉴定为候选的抑癌基因，但是没有发现它们有较高的突变失活频率。这些基因也许可以作为抑癌基因的候选者，因为它们所在的染色体位置经常在癌症中被删除（表 24-3）。具体的例子包括肺癌及其他肿瘤中位于染色质 3p 的 *RASFF1A* 和 *FHIT*（Dammann et al. 2000；Burbee et al. 2001）。其他的则是已知的参与抑制肿瘤发生通路基因，如促凋亡蛋白基因或 DAP 激酶（Katzenellenbogen et al. 1999）。尽管这些基因的启动子在肿瘤中常被高甲基化，但是对于癌症表观遗传学这个领域它们代表了一类重要的问题，因为许多此类基因极少或者根本不发生突变。我们需要去证明这些基因到底怎样导致了肿瘤的发生。随后的章节将会回到这一问题。

通过随机鉴定因启动子高甲基化导致的异常基因沉默发现了第三类基因（Suzuki et al. 2002；Yamashita et al. 2002；Ushijima 2005）。与第二类基因不同的是，这类基因的功能是未知的，所以也就很难确定这些基因参与到癌症发生的哪些方面。

## 4.2 在癌症细胞中寻找新的表观沉默基因

寻找癌细胞中新的表观遗传沉默基因的最常用方法，是鉴定那些候选的抑癌基因是否存在于某些肿瘤中，是否表达量很低或完全没有。另一种使用到的方法是随机筛选癌基因组中高甲基化的基因（Toyota et al. 1999；Suzui et al. 2002；Yamashita et al. 2002；Ushijima 2005）。这虽然会极大地丰富我们关于癌症生物学的知识，但是也会带来更多挑战。正如最近的综述文章所说，每种方法都有自己的优点和缺点（Ushijima 2005）。

若干方法均起始于以一种或多种根据 CpG 位点甲基化状态对 DNA 进行不同切割的酶的消化。为了鉴定可能的甲基化基因，酶切产物用二维凝胶电泳进行分析，通过随机引物扩增或差减技术区分普通 DNA 和癌症 DNA 序列中的甲基化的部分。这些技术可以高通量地鉴定高甲基化的基因，并都增进了我们对那些重要的候选抑癌基因的认识。虽然，鉴定出的序列可能会也可能不会和 CpG 岛结合在一起，但这些 CpG 岛关键性地定位于参与基因沉默的区域。有时在这些进程的最终产物中常出现高甲基化的重复

序列。因此，高甲基化抑癌基因被鉴定的效率不高，且很难覆盖整个基因组范围。

其他的方法依赖于将基因组的 CpG 岛序列点在微阵列芯片上（Chen et al. 2003），通常会考虑其与基因起始位点的关系，然后用被甲基化敏感的酶消化的 DNA 或 cDNA（如果研究基因表达的话）来作为探针与芯片进行杂交。对高甲基化的抑癌基因这一方法可能非常有效，但会受到可被点在芯片上的 CpG 岛的数量的限制。

另一种方法是用引起 DNA 去甲基化的药物来操作培养的肿瘤细胞，如 5-氮杂胞苷（5-az）或 5-氮杂-2′脱氧胞苷（5-aza-2′-lexytidine），用药物处理前和处理后的细胞 RNA 与基因微阵列芯片杂交从而发现哪些基因表达上调（Suzuki et al. 2002；Yamashita et al. 2002）。这种方法也许可以鉴定出每个可培养的人类癌症细胞中所有的高甲基化基因。然而，药物去甲基化活性以外的导致基因表达变化的活性会降低高甲基化基因的鉴定效率。更少被认识到的是，那些处理前或处理后均低表达的基因对芯片灵敏性的要求是大部分芯片所无法达到的，因此会显著降低该方法的效率（Suzuki et al. 2002）。药物处理后使用差减法富集转录物可以提高基因芯片方法的灵敏度（Suzuki et al. 2002），但是这必须使已能覆盖整个基因组的芯片技术能适用于荧光标记。同时使用那些可以协同启动子 DNA 甲基化改变核小体结构的药物，如组蛋白去乙酰化抑制剂，有助于这种基因芯片依赖的方法。这个策略有助于更加特异性地寻找因高甲基化发生了核小体结构改变的沉默基因，细节将在后续章节详尽讨论。用后一种方法在直肠癌中鉴定了许多重要的沉默基因（Suzuki et al. 2002）。

## 4.3　确认癌症中高甲基化基因的重要作用

在癌症中大量的高甲基化基因的发现对于科学研究来说是一种巨大的挑战。基因启动子的频繁甲基化本身并不像遗传突变那样能保证对该基因进行沉默。当某些高甲基化基因并非已知的经典抑癌基因，也无证据表明这些基因在肿瘤细胞中会经常发生突变时尤其如此。因此，该被研究基因必须通过其编码蛋白的性质和其与癌症的相关性来确认该基因表达缺失在功能上的重要性。这样的研究分为许多不同阶段，每一步都进一步增强其在癌症发生中的作用（表 24-4）。

①是关于高甲基化和它对基因表达状态的影响记录，包括伴随基因启动子去甲基化基因的再表达能力。②高甲基化和基因沉默的关系必须在原代和经培养的肿瘤细胞中都被确认。③需要知道基因沉默发生在肿瘤发生的哪一阶段（图 24-4）。④基因的功能缺失对于肿瘤发生的贡献必须被直接证明。这可以始于常规的细胞培养以评估重新导入基因对细胞功能的影响，如细胞凋亡的诱导，对培养基上克隆形成的影响以及在裸鼠中异种移植细胞对肿瘤形成的影响。⑤编码蛋白的功能必须被建立，或者通过以往的此类相关蛋白的知识，或者通过由蛋白质结构暗示的某种功能，或者通过在细胞培养模型中研究该蛋白的生物学功能。⑥最终必须实现用转基因敲除技术去确认该基因具有抑癌基因功能并了解编码蛋白在发育以及成体细胞更新中的作用。通过对癌细胞整个基因组范围的筛选发生杂合性丢失（LOH）基因发现了 *HIC-1* 基因，之后对基因敲除小鼠的研究直接证明了 *HIC-1* 具有抑癌基因的功能（W. Y. Chen et al. 2003，2004）。这些挑战，尤其是第六步，揭示了在癌症中发现表观遗传沉默基因的价值，但为本领域研究者提供

了大量的的工作任务。

**表 24-4 鉴别对肿瘤发生重要的高甲基化基因的步骤**

1. 基因启动子 CpG 岛甲基化并和其转录沉默相关，这种沉默在培养的细胞中可以被去甲基化酶药物所打破。
2. 在肿瘤细胞（培养的细胞或初代肿瘤）中有特异的启动子高甲基化，而不在正常细胞中发生。初代肿瘤细胞中有明显的高甲基化改变。
3. 检测特定癌症肿瘤发生过程中高甲基化改变的位点。
4. 检测基因重插入对培养的细胞产生的肿瘤发生相关的基因沉默的潜能，以及在琼脂克隆细胞或裸鼠中肿瘤细胞生长等的影响。
5. 建立这个沉默基因的蛋白产物并通过对其活性的检测或其特征结构域的确认。
6. 通过小鼠敲除试验确定肿瘤抑制基因在细胞自我更新等方面的活性和功能，特别是那些完全未知的基因。

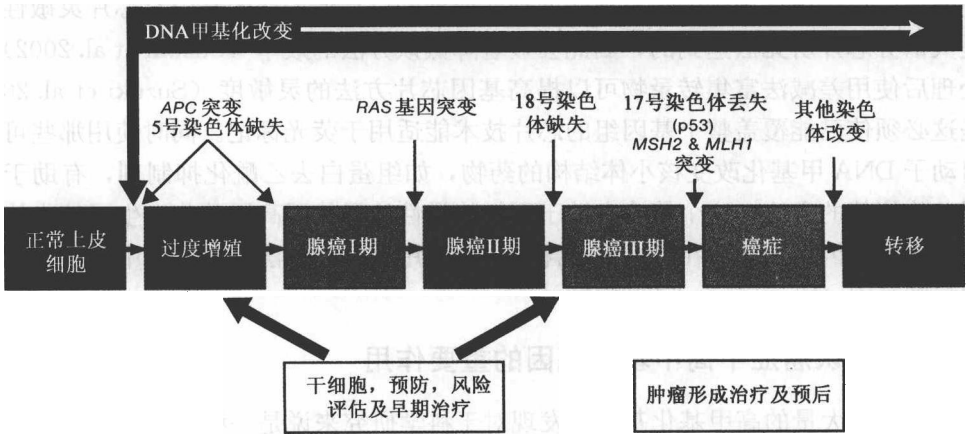



图 24-4 DNA 异常甲基化在肿瘤发生中的早期功能

直肠癌发生过程中遗传改变的经典模型（Kinzler and Vogelstein 1997）。DNA 甲基化改变发生在非常早期（红色箭头所示）阶段，在从普通细胞向增生上皮细胞转变的时期。它战略上处于非常重要的位置（左下黑箭头和左下方框），与重要的遗传变化一起，调控干细胞进入异常克隆扩增状态。如左下黑框所示，这些表观遗传异常也可作为癌症发生的标志。在非侵染性到侵染性，直至最终的肿瘤扩散阶段，DNA 异常甲基化持续增加（右下箭头和右侧方框）。这对癌症治疗和作为诊断标志都具有意义。

## 5. 表观遗传基因沉默及其在癌症形成中的作用——对早期肿瘤形成阶段的重要性

在经典的癌症演化的观点中，如 Vogelstein 及其同事所论述的（Kinzler and Vogelstein 1997），一系列的基因改变导致癌变从早期恶变前阶段，经浸润性癌，最终发展到癌扩散（图 24-4）。这种进程并不一定在所有的癌症中以相同的严格线性顺序发展。我们知道贯穿整个癌症发生事件，表观遗传改变也在同时发生：早期的表现既有普通 DNA 甲基化的大量缺失也有我们讨论过的更多的基因启动子位点高甲基化。因此，有可能表观遗传和遗传事件相互作用导致了细胞的异常并且贯穿了癌症发生的全进程。本章中，关于表观遗传的两方面数据——基因印记缺失（lost of imprinting，LOI）（见

第 23 章）和基因沉默被证明对于癌症发生最早期阶段极为重要。

LOI 涉及在肿瘤形成过程中沉默的印记等位基因激活的进程，因此该基因出现双等位基因表达并产生过量基因产物（Rainier et al. 1993）。研究最为透彻的是在直肠肿瘤中的 *IGF2* 基因（Kaneda and Feinberg 2005）。在这个例子中，位于染色体 11p 上的印记基因 *H19* 启动子高甲基化会导致一种复杂的染色质调控进程（见第 19 章），从而异常地激活沉默的 *IGF2* 基因（Kaneda and Feinberg 2005）。*IGF2* 的双等位基因表达会导致大量的生长因子 IGF2 蛋白的产生。实验证据表明这在直肠癌形成的最早期阶段可能起到一定作用（Kaneda and Feinberg 2005；Sakatani et al. 2005）。事实上，最近的小鼠模型研究暗示单独的 LOI 事件也许就足以起始肿瘤发生（Hlom et al. 2005）。

第二常见的肿瘤发生转化是表观遗传导致的基因沉默，发生在癌症早期阶段。这与细胞受胁迫和暴露于疾病发生状态有紧密联系。涉及的基因常参与受胁迫的细胞在 DNA 损伤或慢性伤害中存活下来，然后像干细胞/前体细胞一样无性扩增，继而倾向于发生后续的表观遗传和遗传事件以促进肿瘤的形成（图 24-5）。第一个证明这类参与的证据来源于一些经典的抑癌基因的数据，这些基因在人类癌症中要么发生突变要么被表观遗传沉默。例如，*p16 ink4a* 的表观遗传沉默很早就出现于恶变前细胞，在肿瘤（如肺癌）形成前的早期变化中（Belinsky et al. 1998）以及某些妇女正常乳房的少量乳腺增生表皮细胞中（Holst et al. 2003）。对于在实验环境中细胞培养液内（培养皿中）生长的正常人类乳腺表皮细胞，*p16* 基因的沉默是细胞转化早期步骤的前提（Kiyono et al. 1998；Romanov et al. 2001）。该基因的功能缺失会导致一部分乳腺细胞不能进入凋亡检查点，当它们继续增殖时，这些细胞的染色质异常进一步增加并表达端粒酶。

第二个有关错配修复基因的例子是 *MLH1* 基因。该基因在易患有某型直肠癌的家系成员的生殖系细胞中有突变，这种癌症伴随着多种遗传变化并被称为微卫星脆性表型（Liu et al. 1995）。然而，10%～15%的非家族性直肠癌患者也具有这种表型，并且这些癌症患者大部分发生非突变 *MLH1* 基因的表观遗传沉默（Herman et al. 1998；Veigl et al. 1998）。在培养的细胞中，重新表达沉默的 *MLH1* 基因会导致有功能的蛋白重新出现从而恢复相当部分的损伤错配修复（Herman et al. 1998）。

最近，一个检查点调节基因 *Chfr* 同时还调控着另一种基因组的完整性，即染色体稳定性和正确的倍性，该蛋白在肿瘤中会发生突变但是更多的是被表观遗传沉默，这发生于肺癌及其他癌症中，尤其的是发生于直肠癌的早期阶段（Mizuno et al. 2002）。小鼠基因敲除研究揭示了其抑癌基因功能，它主要是通过作为一个 E3 泛素连接酶来调节一个有丝分裂调节基因 *Aurora A*。从敲除小鼠获得的胚胎细胞染色体不稳定并且易发生转化。

因为癌症中高甲基化基因的名单越来越长，癌症形成早期的重要沉默事件也被用来定义抑癌基因的候选者，这些基因通常没有突变只是发生表观遗传改变。例如，DNA 修复基因，*O6-MGMT*，在直肠癌形成早期发生沉默（Esteller et al. 2001a），该基因的功能缺失使鸟嘌呤的烷基化更易被保留下来，从而发生 G⟶A 的突变。事实上，该基因的沉默发生在直肠息肉恶化成肿瘤之前，就在稍后的肿瘤形成期 *P53* 和 *RAS* 发生大量突变之前（Esteller et al. 2001a；Wolf et al. 2001）。相似的，把细胞放入一种氧化

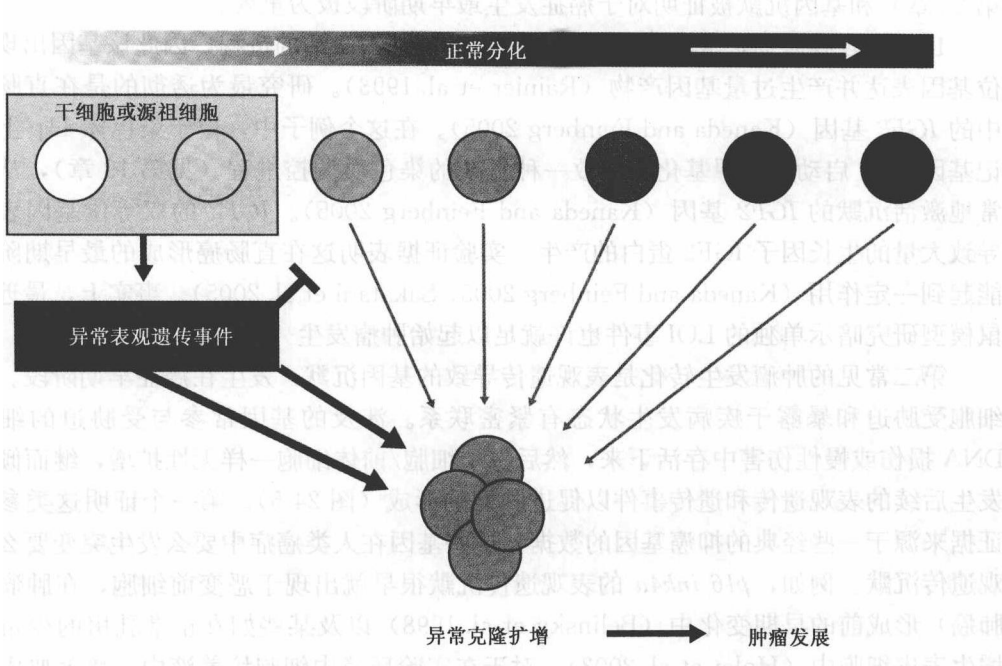



图 24-5 表观遗传基因沉默事件和肿瘤发生

肿瘤发生最早的一步被称为异常的克隆扩增，这缘起细胞自我更新的胁迫。可以由多种因素引起，如衰老、慢性损伤和炎症。这些细胞克隆具有高风险，随后会发生导致肿瘤发生的遗传或表观遗传事件。对于从成体细胞自我更新系统中诱导干细胞或前体细胞形成异常的克隆扩增，在许多情况下，其最早期的可继承因素，可能是异常的表观遗传事件，如本章所关注的异常基因沉默。基因沉默由抑制转录的组蛋白修饰引起，文中所讨论的染色质上的 DNA 高甲基化则作为一个牢固的锁，使这种沉默可以被稳定继承。反过来，基因沉默也破坏正常的细胞内环境，从而阻止干细胞和前体细胞沿正常的分化途径分化为表皮细胞，形成异常的克隆扩增。

破坏腺嘌呤的环境中，在几乎所有会导致前列腺癌的癌变前损伤中，*GST-Pi* 基因都会发生沉默（Lee et al. 1994）。

第三类沉默的基因是通过随机筛选癌细胞基因组中表观遗传沉默的基因时发现的。这同样也对我们理解癌症早期表观遗传沉默的作用作出了重要贡献。一种特殊而有趣的现象发生在直肠癌形成过程中：早期通过芯片方法研究发现，表观遗传性功能缺失发生在一个基因家族中（Suzuki et al. 2002），该基因家族可以早期异常地激活疾病发生发展的通路。分泌卷曲相关蛋白基因（*secreted frizzled related protein*，*SFRP*）（Suzuki et al. 2004）的转录失活消除了 Wnt 配体与其膜受体相互作用的拮抗信号（Finch et al. 1997）。此失活与 Wnt 介导的全细胞水平上的 β-catenin 的上调有关，尤其是因为核内这个转录因子的数量和活性的增加（Suzuki et al. 2004）。这种转录是 Wnt 通路活性增加的主要特征（Morin et al. 1997；Gregorieff and Clevers 2005）。最重要的是，SFRP 沉默发生在可能引起直肠癌的非常早期的损伤中，并且在 Wnt 通路下游蛋白发生突变以前，会导致核内的 β-catenin 激活（Morin et al. 1997；Gregorieff and Clevers 2005）。

因此，表观遗传事件早期激活的 Wnt 通路稳定地保证了细胞的扩增，和更易于激活突变事件所需要的通路。维持表观遗传的改变，通过 Wnt 通路诱导的细胞内 β-catenin 的上调；维持遗传的改变，通过破坏降解 β-catenin 的蛋白复合物和激活 Wnt 突变来实现。这两方面互相补充从而导致疾病的发生（Suzuki et al. 2004）。

此类基因的另一个例子是 *HIC-1*（*hypermethylated-in-canceer 1*），该基因编码一个锌指转录抑制因子。*HIC-1* 基因是通过在癌细胞染色体缺失热点区域中随机筛选高甲基化 CpG 岛而发现的（Wales et al. 1995）。该基因在癌症形成早期阶段就被沉默但没有发生突变，它在小鼠敲除模型中被证明是抑癌基因（W. Y. Chen et al. 2003，2004）。它通过功能缺失部分互补 *p53* 突变体，上调 SIRT1 的表达（Chen et al. 2005）。SIRT1 是一种重要的感受细胞胁迫并促进干细胞/前体细胞生长的蛋白质（Howitz et al. 2003；Nemoto et al. 2004；Kuzmichev et al. 2005）。

根据以上讨论的数据得出了图 24-4 所描绘的主要假说。这暗示了在癌症演化中某些最早发生的可继承变化可能是表观遗传变化，这通常涉及通过启动子 DNA 甲基化来维持的严格的基因转录水平沉默。进一步理解这些情况所遇到的挑战是要把在癌症中研究表观遗传变化遇到的问题整合起来（表 24-5），更多讨论将在下文进行。解决这些挑战，尤其是理解表观遗传变化对肿瘤发生最早期的贡献，可能会极大地丰富预防或早期干预癌症发生的分子生物学手段。

**表 24-5　研究癌症表观遗传基因沉默分子机制的主要问题**

1. 同一个肿瘤细胞中 DNA 甲基化的获得和丢失同时发生，它们之间的联系有待阐明。
2. 那些将转录活性区域和沉默区域分开的边界元件的分子本质以及它们在肿瘤发生过程中的变化是怎样的？它们怎么阻止异染色质的蔓延？是关键的组蛋白修饰，还是绝缘子结构，或是染色质重建蛋白担起了这项重任。
3. 癌症进程中组蛋白修饰、DNA 甲基化及基因沉默的顺序是怎样的？哪种酶或蛋白复合体是首当其冲的？
4. 对那些稳定的基因沉默来说，哪些甲基化酶对从头甲基化和甲基化维持是特异的？哪些蛋白复合体参与其中或和组蛋白密码的关键组分工作？
5. 染色质和 DNA 甲基化机制的全景是怎样的？它们各部分的对沉默的重要性如何？它们怎么样从沉默中恢复？

## 6. 分子水平剖析表观遗传沉默的癌症基因

在癌细胞中沉默的基因对于我们理解癌症的发生和维持是非常重要的。它们也是一些很好的模型，使我们能理解基因沉默是如何建立和维持的以及哺乳动物基因组是如何将转录区域包装起来并抑制转录的。本书多个章节所强调的对染色质功能的理解，将促进我们理解对是什么启动了癌症中基因的异常沉默以及这些沉默的组件是如何维持其转录抑制的。

许多实验室的工作对于我们理解在癌症细胞中沉默基因启动子的高甲基化 CpG 岛附近的染色质状态做出了很大的贡献。这些研究同时也强调了这些基因周围的染色质与在其具有本底表达水平时有什么差异。在正常细胞中或癌症细胞中转录没有被抑制的基因，都在未甲基化的 CpG 岛内有一个染色质开放区域，其核小体具有不规则的间距，可以检测到一些核酸酶超敏感位点，并且其重要的组蛋白位点会发生活跃基因特异性的

翻译后修饰。激活性组蛋白共价修饰包括 H3K9 和 K14 的乙酰化（H3K9ac 和 H3K14ac）以及 H3K4 的甲基化（Nguyen et al. 2001；Fahrner et al. 2002）。

在以上所述的染色质开放区域的 5′和 3′端，染色质结构上有显著的转变，即 CpG 岛两侧为具转录抑制特征的基因组区域（图 24-6）。在这些边界区域，有一些频率较低的 CpG 位点被甲基化，并且会募集一些甲基化胞嘧啶结合蛋白（MBD）和它们的互作蛋白（如组蛋白去乙酰化酶）到甲基化的 CpG 位点（第 18 章）。因此，CpG 岛以外的区域表现为更易被与转录沉默相关的组蛋白甲基转移酶所接近。所有这些因子导致的结果就是，组蛋白上重要的残基去乙酰化，以及与转录抑制相关的组蛋白甲基标记出现，尤其是 H3K9 二甲基化（Nguyen et al. 2001；Fahrner et al. 2002；Kondo et al. 2003）

这些活跃型和抑制型染色质模式区域交错存在（图 24-6）暗示含有活跃基因启动

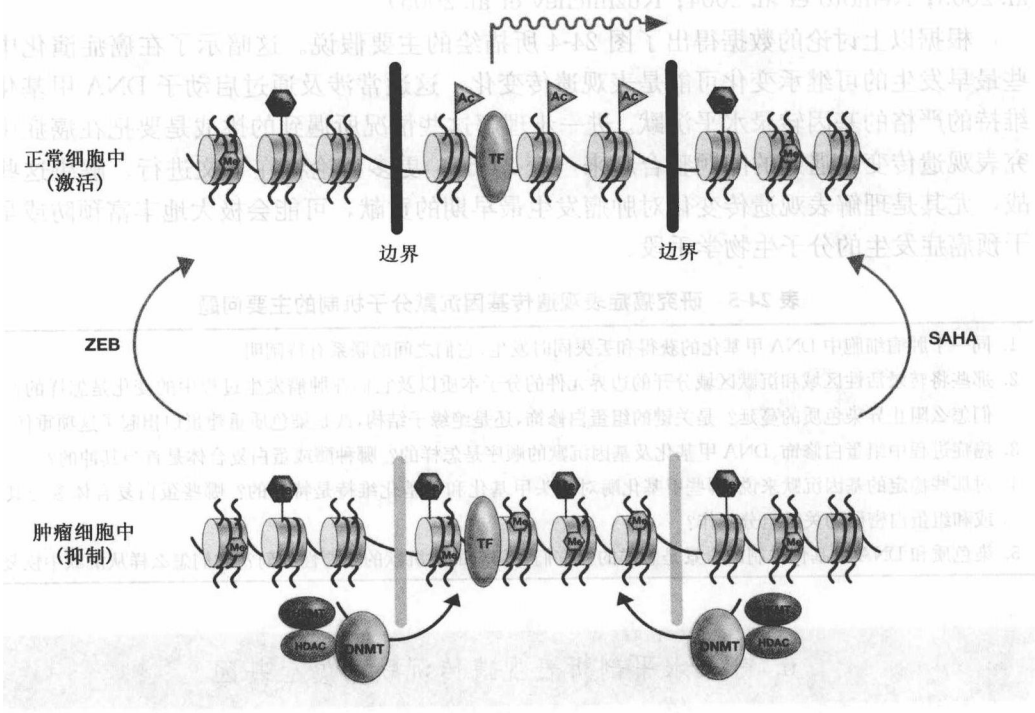



图 24-6 正常和肿瘤细胞中基因启动子 CpG 岛区域内的 DNA 甲基化和组蛋白修饰关系的模型（上图）在表达基因中，边界序列的分子机制还不清楚，但是它可以保护转录起始位点（绿箭头）周围的 CpG 岛不被甲基化。与被保护区域相比，其周围的 CpG 位点 DNA 被甲基化（标有 M 的粉红色六边形），并且具有重要的沉默标记如 H3K9 的甲基化（标有 Me 的红色六边形）。在被保护区域内，重要的组蛋白末端氨基酸（如 H3K9）处于乙酰化状态（标有 Ac 的小旗），转录因子（标有 TF 的椭圆形）结合到转录起始位点区域。（下图）当同样的基因在癌细胞中被异常沉默时，由于边界序列的消失，CpG 岛 DNA 发生高甲基化。这些甲基化由 DNA 甲基转移酶复合物（DNMT）、含有去乙酰化酶（HDAC）的甲基化胞嘧啶结合蛋白复合物和催化组蛋白末端基因沉默标记如 H3K9 甲基化的组蛋白甲基化转移酶（HKMT）共同维持。转录因子不再有活性（箭头消失）。目前癌症主要的表观遗传临床试验方法是通过 DNA 甲基化转移酶抑制剂抑制 DNA 高甲基化或组蛋白去乙酰化酶抑制剂以恢复重要组蛋白氨基酸位点的乙酰化状态。如文中所讨论的，某些最有前景的癌症治疗方法是共同使用 DNMT1 和 HDAC 的抑制剂。

子的 CpG 岛位于某种被“保护”的区域，或者说不被染色质抑制性标记和 DNA 甲基化所靶定（Nguyen et al. 2001；Fahrner et al. 2002；Kondo et al. 2003）。与这些概念类似的是分子“边界”存在于表达基因 CpG 岛启动子的 5′和 3′分界线上。一个主要的问题是准确定义这些边界的本质。目前，候选者是标记了启动子被保护区域的组蛋白修饰本身，直接支持转录激活的转录激活因子和共激活因子复合物，完成核小体置入和（或）移动，从而可能将基因为标记为活跃状态的蛋白质复合物（如核小体重塑复合物）。这些可能会促进转录激活复合物的接近，以支持转录激活的组蛋白变体如 H3.3（第 13 章）替换常规组蛋白，以及绝缘子蛋白复合物及其识别序列的行为。以上一个或多个候选进程是如何使活跃基因启动子所处的未甲基化 CpG 岛周边染色质维持于有利转录的环境的研究，使得癌症中沉默的基因成为我们理解哺乳动物基因组中基因的表达调控的良好模型。

在肿瘤发生过程中，启动子 CpG 岛的转录激活型染色质结构被转变的方式仍然有许多问题，对这种转变的解析是癌症表观遗传领域最重要的挑战之一。如前文所述，异常的基因启动子 CpG 岛甲基化及其伴随的基因沉默，通常出现在癌症发生的非常早期阶段和恶变前阶段。使得理解这些因子对于我们关于癌症生物学的认识可能具有非常重要的意义。这同时也提供了新的可能性以评估癌症风险因子，促进癌症早期诊断以及考虑新的癌症预防和早期干预措施。

本章后文还将继续讨论这一问题，无论涉及怎样的精细机制，最终的结果是在肿瘤发生过程中，用于维持启动子 CpG 岛区域免遭抑制性染色质从 5′和 3′边界侵入的机制“失效”了。结果就将启动子的转录开放模式转变为一种关闭模式，它有紧密的核小体并且失去了核酸酶超敏位点，还出现组蛋白去乙酰化和甲基化等多种转录抑制性染色质的特征（图 24-6）。所有这些事件的等级，或者说哪个事件是维持这种抑制的最重要因素，仍有待确定。然而，通过多基因分析发现的最重要的特征是，无论 DNA 甲基化在起始和（或）维持基因沉默中起的作用是什么，一旦被建立，它对于转录抑制状态的继承性起着主导性作用。因此，通过专一性药物来抑制组蛋白去乙酰化并不能导致那些启动子 CpG 岛高甲基化的基因重新被表达。然而，先用少量去甲基化药物如 5′-脱氧氮杂胞苷（DAC）处理，则组蛋白去乙酰化酶具有累加性或协同性作用，引起沉默基因的重新表达（Cameron et al. 1998；Suzuki et al. 2002）。这个例子，结合芯片分析寻找重新表达的沉默基因，已经证明对寻找癌症中新的表观遗传沉默基因非常有用（Suzuki et al. 2002）。同样的，伴随着 DAC 处理的基因重新表达会导致表观遗传沉默的癌基因启动子区域内抑制性组蛋白甲基化标记的去除（Nguyen et al. 2001；Fahrner et al. 2002；Kondo et al. 2003）。在这个意义上，DNA 甲基化可能作为一个“锁”来稳定癌症基因的表观遗传沉默并保证基因功能性缺失被稳定继承（图 24-6）。

另一个在癌症基因的表观遗传沉默领域内的重要问题是哪些 DNA 甲基化酶负责建立和维持与这个进程相关的启动子 DNA 异常甲基化。目前的数据暗示这些步骤可能是通过与经典的在小鼠发育研究中建立的不同的 Dnmt 的作用而实现的。在小鼠发育中，通过研究三种已知具有生物学活性的 DNA 甲基化转移酶的基因敲除小鼠得知：Dnmt3a 以及 Dnmt3b 负责甲基化的建立（从头 DNA 甲基化），而 Dnmt1 则负责维持已建

立的模式（维持性 DNA 甲基化）（Li et al. 1992；详见第 18 章）。然而，直肠癌细胞培养中，破坏 *DNMT* 基因的研究暗示总体 DNA 甲基化主要部分的维持，包括启动子高甲基化及其引起的基因沉默，同时需要 DNMT1 和 DNMT3b 的参与（Rhee et al. 2000，2002）。对其他癌症细胞的研究产生了更多不同的结果，从这两种酶具有某种程度的相互作用以维持甲基化，但 DNMT1 单独下调就会导致基因去甲基化并且重新激活转录（Leu et al. 2003）。因此，关于 DNMT 如何在癌症细胞中建立和维持异常的 DNA 甲基化模式仍需要进一步研究。特别是，我们需要澄清这些酶与哪些复合物相互作用以及它们的基因启动子靶向性的机制。最近许多的研究暗示转录抑制复合物对于这种募集是至关重要的（Di Corce et al. 2002；Fuks et al. 2003；Brenner et al. 2005）。无论什么机制参与其中，我们必须记住，实验上，哺乳动物的 *DNMT* 具有复杂的功能，其不仅包含来自 C 端的 DNA 甲基化活性，还同时带有来自 N 端的直接转录抑制活性（见图 18-3）（Robertson et al. 2000；Rountree et al. 2000；Fuks et al. 2001）。这些酶也可以结合或者潜在的募集重要的转录抑制因子，包括组蛋白去乙酰化酶和甲基化胞嘧啶结合蛋白（MBD）。因此，*DNMT* 的功能在转录沉默中可能具有多面性，从起始到维持，可能参与也可能不参与 DNA 甲基化。癌症中的基因沉默对于解决这些可能性将是一种重要的模式。

## 7. 概述癌症中表观遗传基因沉默的主要研究方向

从以上讨论中，我们已经了解了很多在肿瘤发生过程中导致启动子 DNA 甲基化和基因沉默的分子事件，但是还有更多的事件有待阐明。表 24-5 总结了未来研究必须解决的，至少是一部分最重要问题。①决定整体水平 DNA 低甲基化和许多局部启动子 DNA 高甲基化的分子事件。这种共存状态暗示在癌症细胞中存在染色质的一种广泛的靶向错误。对此的病理学研究应该会指明哺乳动物细胞是如何为正确的基因表达模式进行基因组包装，并如何维持染色体的完整性。②各基因启动子两侧染色质边界的决定因子及其与正常和异常转录状态的关系都需要得以描述。绝缘子是否参与保护正常活性基因不受转录抑制性染色质的侵入？组蛋白修饰本身是否可以建立起边界？这些决定因子在肿瘤发生过程中是如何被改变的？③在癌症细胞中导致基因异常沉默发生的事件顺序是什么？这是否为起始下调基因转录所必需？这些重要的组蛋白修饰，包括组蛋白氨基酸上的去乙酰化和甲基化以及募集介导这些过程的酶（见第 10 章），是否先于并导致 DNA 甲基化？如果是这样，导致这种靶向性的分子互作是什么，以及对正常细胞和癌症干细胞非常重要的沉默复合物，如 Polycomb 蛋白家族（见第 11 章）是如何起作用的？④确定到底哪些 DNMT 是癌症中启动子 DNA 甲基化起始和维持所需要的，以及它们是如何互作的？以上讨论的哪些染色质成分可能与这些酶互作并且定位它们？⑤一旦异常的可继承基因沉默在癌症细胞中被建立，维持它的分子步骤的具体等级是什么？最后的这个问题不仅仅是一个重要的基础研究问题，同时也在应用研究上，是癌症预防和治疗中恢复表达沉默基因的核心内容。

## 8. 通过 DNA 甲基化检测癌症

点状 CpG 岛高甲基化情况在癌症细胞中非常普遍，再加上有高灵敏度的检测甲基化的能力，导致发展出了许多种在体液中检测癌症的方法。亚硫酸氢钠处理可以将 DNA 链上的胞嘧啶转化为尿嘧啶而 5-甲基胞嘧啶则不会发生改变，从而能从正常细胞的背景中检测到获得性的 CpG 岛甲基化变化。PCR 方法如甲基化特异性 PCR（MSP），其引物被设计成只扩增甲基化的区域，非常灵敏。别的方法包括以实时 PCR 为基础的技术如“甲基光”，使用只只结合甲基化 DNA 的荧光探针来检测甲基化模式（见本章题图左，以基因为 Y 轴，肿瘤为 X 轴）。这些技术可以在 1000～10 000 个基因的背景下检测到一个甲基化的基因。因此，异常的甲基化模式可以很容易被检测到；这些方法可以应用于细胞的混合物甚至多种体液混合物如血浆、尿液或者痰液（Laird 2003）。通过鉴定胞嘧啶甲基化的变化来检测癌症稳定性非常好，因为 DNA 本身的稳定性较 RNA 和蛋白质更高。同时，也因为甲基化模式的改变通常是癌症类型特异的，这些方法也许可以区分不同的癌症。如前所述，现在许多研究提供了“原则性证据”，使利用高甲基化 DNA 启动子序列成为非常有前途的癌症预测和检测手段。这种方法的医学应用价值还需要在更大规模的研究中证实。此类研究必然会在未来几年中展开。

## 9. 表观遗传治疗

通过 DNA 甲基化和核小体修饰的改变导致的癌症相关基因沉默使人们认识到沉默的核小体可能也是可以用于治疗的靶点。因此，一种称为“表观遗传治疗法”的新方法被发展出来，单独或混合使用可改变核小体或 DNA 甲基化模式的药物，可以实现对治疗效果的影响（Egger et al. 2004）。

核苷类似物，5-氮杂胞苷、5-氮杂脱氧胞苷和泽布拉另（Zebularine）都是非常有效的胞嘧啶甲基化抑制剂（图 24-7a）。当它们被代谢为相应的脱氧核苷三磷酸后，会在复制的细胞中进入 DNA。一旦被装入到 DNA 中，它们就会与所有已知的三种 DNA 甲基转移酶相互作用形成共价中间体，最终在下一轮的 DNA 合成中抑制 DNA 甲基化。这些化学药物的作用机理已经被了解得非常透彻，并且已经在组织培养或移植瘤模型中使用了相当长的一段时间，用于激活沉默的基因。然而，最近发现他们可以被用于某种血液肿瘤的治疗，尤其是骨髓增生异常综合征（MDS），这是一种白血病的早期状态，主要发生在年老的患者中（Lubbert 2000；Wijermans et al. 2000；Silverman et al. 2002；Issa et al. 2004）。对这些综合征，以及从白血病前期发展过来的白血病患者的临床效果日益显著。相应的，5-氮杂胞苷和 5-氮杂脱氧胞苷药物被临床命名为维达扎（Vidaza）和德西他滨（Decitabine），分别已得到美国食品和药物管理局（FDA）的批准，被用于治疗这些患者。泽布拉另（Zebularine），也是一种 DNA 甲基化转移酶抑制剂，它正处于临床研发的早期阶段。目前，尚未发现无需整合入 DNA 的有效抑制剂，但这可能是临床上更需要的，因为它们的副作用可能更小。许多方法正在进行以合成或发现这样

的药物。

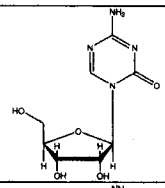
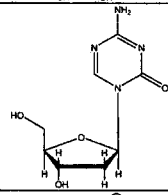
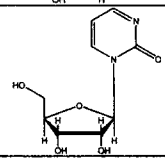
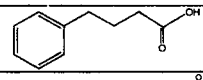
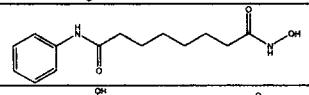
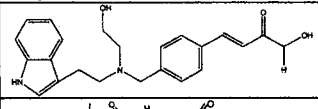
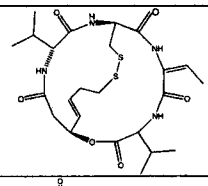
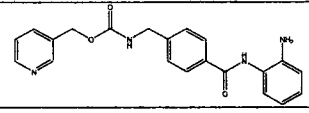
| | 成分 | 结构 | 癌类型 | 临床期 |
|---|---|---|---|---|
| a | DNA 甲基化抑制剂 | | | |
| | 5-Azacytidine<br>5-Aza-CR<br>Vidaza |  | MDS;<br>Hematologic malignancies | Ⅰ，Ⅱ和Ⅲ已经FDA批准用于MDS |
| | 5-Aza-2'-deoxycytidine<br>5-Aza-CdR<br>Dacogen |  | MDS;<br>Hematologic malignancies | Ⅰ，Ⅱ和Ⅲ |
| | Zebularine<br>1-β-D-ribofuranosyl-2(1H)-pyrimidinone |  | N/A | 研发早期 |
| b | 组蛋白去乙酰化抑制剂 | | | |
| | 4-Phenylbutyrate (PBA) |  | Refractory solid tumors | Ⅰ |
| | Suberoylanilide hydroxamic acid (SAHA) |  | Solid tumors and hematologic malignancies | Ⅰ，Ⅱ |
| | NVP-LAQ824 |  | N/A | Ⅰ |
| | Depsipeptide<br>FK-228<br>FR901228 |  | Advanced neoplasms, CLL, AML, and T-cell lymphoma | Ⅰ，Ⅱ |
| | MS-275 |  | 实体瘤和淋巴瘤 | Ⅰ，Ⅱ |

图 24-7 DNA 甲基化核苷类似物抑制剂（a），和组蛋白去乙酰化抑制剂（b）的结构

(a) 已知的三种能在整合入 DNA 后抑制 DNA 甲基化的核苷类似物。5-az（Vidaza）和 5-aza-deoxycytidine (Decitabine) 已被批准用于白血病的治疗。泽布拉另处于开发的早期阶段。(b) 众多组蛋白抑制剂的一些例子，其中有些处于临床试验阶段（译者注：SAHA 已经被 FDA 批准）。

尽管维达扎（Vidaza）和德西他滨（Decitabine）已经显示了很强的临床效果，但是还很难明确的证实这些药物的靶位点就是那些甲基化的基因启动子。初步的实验表明经过 Decitabine 处理后 *p15* 抑癌基因发生了去甲基化（Daskalakis et al. 2002）；然而，还需要证明这些药物的疗效是由诱导基因表达还是其他机制造成的。

使用组蛋白去乙酰化酶抑制剂的临床试验也在进行（图 24-7b）。已知多种药物可以起到显著的组蛋白去乙酰化酶抑制作用。其中某些，如苯丁酸或丙戊酸，已经在临床中被运用了一段时间，但被用于处理其他情况（Marks et al. 2001；Richon and O'Brien 2002）。这些药物抑制所有的去乙酰化酶，而且我们也不清楚它们对特定肿瘤的疗效是否只是因其对组蛋白去乙酰化的抑制而非某种其他反应。新的化合物，如 SAHA 和酯肽类物质，是更专一的组蛋白去乙酰化酶抑制剂，也展示了良好的临床效果（图24-8）。然而，仍然很难确定这些药物在患者身上的疗效是否是由其理论机制所造成。

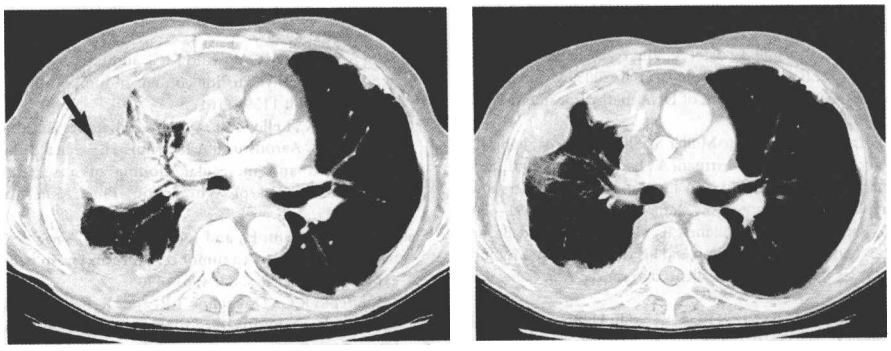


图 24-8　组蛋白去乙酰化酶抑制剂可以引起肿瘤消退

间皮瘤患者的计算机断层扫描图像。（左），在治疗前右肺内有一大型分叉状肿块（箭头所示）挤压正常肺（黑色）。（右），经 SAHA 治疗四个月，肿瘤缩小且右肺明显扩增（由 Paul Marks 博士提供，Memorial Sloan-Kettering Cancer Center，NY）。

以前文讨论的 DNA 甲基化和核小体的相互作用为基础的新的临床试验，聚焦于将 DNA 甲基化抑制剂和组蛋白去乙酰化抑制剂结合使用，以图用更小剂量的药物来获得更好的疗效。为了验证这种假说，一些令人激动的临床试验正在进行当中；然而，还不清楚这些方法是否会有效。

虽然表观遗传治疗方法有良好的理论基础，但某些药物缺乏专一性也导致了一些担忧。例如，核苷类似物就是非特异性的 DNA 甲基转移酶抑制剂并且在全基因组上抑制 DNA 甲基化。因此，治疗可能会导致非特异的基因激活，尽管这在那些接受治疗的重病患身上看起来不是主要问题。类似的，除了去乙酰化以外还有一些其他的核小体修饰，如前面各章所讨论的组蛋白氨基酸上的甲基化，这些都可以作为潜在的治疗靶点。因此，人们有极大的兴趣去寻找能激活沉默基因的新的治疗靶点，并且可以预期在未来数年将会有令人激动的进展。

袁　文　译
谢精卫　校
朱　冰　审校

# 参考文献


Amann J.M., Nip J., Strom D.K., Lutterbach B., Harada H., Lenny N., Downing J.R., Meyers S., and Hiebert S.W. 2001. ETO, a target of t(8;21) in acute leukemia, makes distinct contacts with multiple histone deacetylases and binds mSin3A through its oligomerization domain. *Mol. Cell. Biol.* **21:** 6470–6483.

Baylin S.B., Fearon E.R., Vogelstein B., de Bustros A., Sharkis S.J., Burke P.J., Staal S.P., and Nelkin B.D. 1987. Hypermethylation of the 5′ region of the calcitonin gene is a property of human lymphoid and acute myeloid malignancies. *Blood* **70:** 412–417.

Belinsky S.A., Nikula K.J., Palmisano W.A., Michels R., Saccomanno G., Gabrielson E., Baylin S.B., and Herman J.G. 1998. Aberrant methylation of $p16^{INK4a}$ is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc. Natl. Acad. Sci.* **95:** 11891–11896.

Brenner C., Deplus R., Didelot C., Loriot A., Vire E., De Smet C., Gutierrez A., Danovi D., Bernard D., Boon T., et al. 2005. Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO J.* **24:** 336–346.

Burbee D.G., Forgacs E., Zochbauer-Muller S., Shivakumar L., Fong K., Gao B., Randle D., Kondo M., Virmani A., Bader S., et al. 2001. Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression. *J. Natl. Cancer Inst.* **93:** 691–699.

Cameron E.E., Bachman K.E., Myohanen S., Herman J.G., and Baylin S.B. 1999. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat. Genet.* **21:** 103–107.

Chen C.M., Chen H.L., Hsiau T.H., Hsiau A.H., Shi H., Brock G.J., Wei S.H., Caldwell C.W., Yan P.S., and Huang T.H. 2003. Methylation target array for rapid analysis of CpG island hypermethylation in multiple tissue genomes. *Am. J. Pathol.* **163:** 37–45.

Chen W., Cooper T.K., Zahnow C.A., Overholtzer M., Zhao Z., Ladanyi M., Karp J.E., Gokgoz N., Wunder J.S., Andrulis I.L., et al. 2004. Epigenetic and genetic loss of *Hic1* function accentuates the role of *p53* in tumorigenesis. *Cancer Cell* **6:** 387–398.

Chen W.Y., Wang D.H., Yen R.C., Luo J., Gu W., and Baylin S.B. 2005. Tumor suppressor HIC1 directly regulates SIRT1 to modulate p53-dependent DNA-damage responses. *Cell* **123:** 437–448.

Chen W.Y., Zeng X., Carter M.G., Morrell C.N., Chiu Yen R.W., Esteller M., Watkins D.N., Herman J.G., Mankowski J.L., and Baylin S.B. 2003. Heterozygous disruption of *Hic1* predisposes mice to a gender-dependent spectrum of malignant tumors. *Nat. Genet.* **33:** 197–202.

Dammann R., Li C., Yoon J.H., Chin P.L., Bates S., and Pfeifer G.P. 2000. Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. *Nat. Genet.* **25:** 315–319.

Daskalakis M., Nguyen T.T., Nguyen C., Guldberg P., Kohler G., Wijermans P., Jones P.A., and Lubbert M. 2002. Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2′-deoxycytidine (decitabine) treatment. *Blood* **100:** 2957–2964.

Di Croce L., Raker V.A., Corsaro M., Fazi F., Fanelli M., Faretta M., Fuks F., Lo Coco F., Kouzarides T., Nervi C., et al. 2002. Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* **295:** 1079–1082.

Egger G., Liang G., Aparicio A., and Jones P.A. 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429:** 457–463.

Ehrlich M., Gama-Sosa M.A., Huang L.H., Midgett R.M., Kuo K.C., McCune R.A., and Gehrke C. 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* **10:** 2709–2721.

Esteller M., Risques R.A., Toyota M., Capella G., Moreno V., Peinado M.A., Baylin S.B., and Herman J.G. 2001a. Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis. *Cancer Res.* **61:** 4689–4692.

Esteller M., Fraga M.F., Guo M., Garcia-Foncillas J., Hedenfalk I., Godwin A.K., Trojan J., Vaurs-Barriere C., Bignon Y.J., Ramus S., et al. 2001b. DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum. Mol. Genet.* **10:** 3001–3007.

Fahrner J.A., Eguchi S., Herman J.G., and Baylin S.B. 2002. Dependence of histone modifications and gene expression on DNA hypermethylation in cancer. *Cancer Res.* **62:** 7213–7218.

Feinberg A.P. and Vogelstein B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301:** 89–92.

Feinberg A.P., Gehrke C.W., Kuo K.C., and Ehrlich M. 1988. Reduced genomic 5-methylcytosine content in human colonic neoplasia. *Cancer Res.* **48:** 1159–1161.

Finch P.W., He X., Kelley M.J., Uren A., Schaudies R.P., Popescu N.C., Rudikoff S., Aaronson S.A., Varmus H.E., and Rubin J.S. 1997. Purification and molecular cloning of a secreted, Frizzled-related antagonist of Wnt action. *Proc. Natl. Acad. Sci.* **94:** 6770–6775.

Flatau E., Bogenmann E., and Jones P.A. 1983. Variable 5-methylcytosine levels in human tumor cell lines and fresh pediatric tumor explants. *Cancer Res.* **43:** 4901–4905.

Frost P., Liteplo R.G., Donaghue T.P., and Kerbel R.S. 1984. Selection of strongly immunogenic "tum" variants from tumors at high frequency using 5-azacytidine. *J. Exp. Med.* **159:** 1491–1501.

Fuks F., Hurd P.J., Deplus R., and Kouzarides T. 2003. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* **31:** 2305–2312.

Fuks F., Burgers W.A., Godin N., Kasai M., and Kouzarides T. 2001. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J.* **20:** 2536–2544.

Gaudet F., Hodgson J.G., Eden A., Jackson-Grusby L., Dausman J., Gray J.W., Leonhardt H., and Jaenisch R. 2003. Induction of tumors in mice by genomic hypomethylation. *Science* **300:** 489–492.

Gazzoli I., Loda M., Garber J., Syngal S., and Kolodner R.D. 2002. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Res.* **62:** 3925–3928.

Grady W.M., Willis J., Guilford P.J., Dunbier A.K., Toro T.T., Lynch H., Wiesner G., Ferguson K., Eng C., Park J.G., et al. 2000. Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. *Nat. Genet.* **26:** 16–17.

Greenblatt M.S., Bennett W.P., Hollstein M., and Harris C.C. 1994. Mutations in the p53 tumor suppressor gene: Clues to cancer etiology and molecular pathogenesis. *Cancer Res.* **54:** 4855–4878.

Gregorieff A. and Clevers H. 2005. Wnt signaling in the intestinal epithelium: From endoderm to cancer. *Genes Dev.* **19:** 877–890.

Hanahan D. and Weinberg R.A. 2000. The hallmarks of cancer. *Cell* **100:** 57–70.

Herman J.G. and Baylin S.B. 2003. Gene silencing in cancer in association with promoter hypermethylation. *N. Engl. J. Med.* **349:** 2042–2054.

Herman J.G., Umar A., Polyak K., Graff J.R., Ahuja N., Issa J.P., Markowitz S., Willson J.K., Hamilton S.R., Kinzler K.W., et al. 1998. Incidence and functional consequences of *hMLH1* promoter hypermethylation in colorectal carcinoma. *Proc. Natl. Acad. Sci.* **95:** 6870–6875.

Holm T.M., Jackson-Grusby L., Brambrink T., Yamada Y., Rideout W.M., III, and Jaenisch R. 2005. Global loss of imprinting leads to widespread tumorigenesis in adult mice. *Cancer Cell* **8:** 275–285.

Holst C.R., Nuovo G.J., Esteller M., Chew K., Baylin S.B., Herman J.G.,

and Tlsty T.D. 2003. Methylation of *p16$^{INK4a}$* promoters occurs in vivo in histologically normal human mammary epithelia. *Cancer Res.* **63:** 1596–1601.

Howitz K.T., Bitterman K.J., Cohen H.Y., Lamming D.W., Lavu S., Wood J.G., Zipkin R.E., Chung P., Kisielewski A., Zhang L.L., et al. 2003. Small molecule activators of sirtuins extend *Saccharomyces cerevisiae* lifespan. *Nature* **425:** 191–196.

Issa J.P., Ottaviano Y.L., Celano P., Hamilton S.R., Davidson N.E., and Baylin S.B. 1994. Methylation of the oestrogen receptor CpG island links ageing and neoplasia in human colon. *Nat. Genet.* **7:** 536–540.

Issa J.P., Garcia-Manero G., Giles F.J., Mannari R., Thomas D., Faderl S., Bayar E., Lyons J., Rosenfeld C.S., Cortes J., and Kantarjian H.M. 2004. Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2′-deoxycytidine (decitabine) in hematopoietic malignancies. *Blood* **103:** 1635–1640.

Jones P.A. 1999. The DNA methylation paradox. *Trends Genet.* **15:** 34–37.

Jones P.A. and Baylin S.B. 2002. The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* **3:** 415–428.

Jones P.A. and Laird P.W. 1999. Cancer epigenetics comes of age. *Nat. Genet.* **21:** 163–167.

Jones P.A. and Taylor S.M. 1980. Cellular differentiation, cytidine analogs and DNA methylation. *Cell* **20:** 85–93.

Kaminskas E., Farrell A., Abraham S., Baird A., Hsieh L.-S., Lee S.-L., Leighton J.K., Patel H., Rahman A., Sridhara R., et al. 2005. Approval summary: Azacitidine for treatment of myelodysplastic syndrome subtypes. *Clin. Cancer Res.* **11:** 3604–3608.

Kaneda A. and Feinberg A.P. 2005. Loss of imprinting of *IGF2*: A common epigenetic modifier of intestinal tumor risk. *Cancer Res.* **65:** 11236–11240.

Katzenellenbogen R.A., Baylin S.B., and Herman J.G. 1999. Hypermethylation of the DAP-kinase CpG island is a common alteration in B-cell malignancies. *Blood* **93:** 4347–4353.

Kinzler K.W. and Vogelstein B. 1997. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* **386:** 761–763.

Kiyono T., Foster S.A., Koop J.I., McDougall J.K., Galloway D.A., and Klingelhutz A.J. 1998. Both Rb/p16$^{INK4a}$ inactivation and telomerase activity are required to immortalize human epithelial cells. *Nature* **396:** 84–88.

Knudson A.G. 2001. Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer* **1:** 157–162.

Kondo Y., Shen L., and Issa J.P. 2003. Critical role of histone methylation in tumor suppressor gene silencing in colorectal cancer. *Mol. Cell. Biol.* **23:** 206–215.

Kuzmichev A., Margueron R., Vaquero A., Preissner T.S., Scher M., Kirmizis A., Ouyang X., Brockdorff N., Abate-Shen C., Farnham P., and Reinberg D. 2005. Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. *Proc. Natl. Acad. Sci.* **102:** 1859–1864.

Laird P.W. 2003. The power and the promise of DNA methylation markers. *Nat. Rev. Cancer* **3:** 253–266.

Laird P.W., Jackson-Grusby L., Fazeli A., Dickinson S.L., Jung W.E., Li E., Weinberg R.A., and Jaenisch R. 1995. Suppression of intestinal neoplasia by DNA hypomethylation. *Cell* **81:** 197–205.

Lapeyre J.N. and Becker F.F. 1979. 5-Methylcytosine content of nuclear DNA during chemical hepatocarcinogenesis and in carcinomas which result. *Biochem. Biophys. Res. Commun.* **87:** 698–705.

Lee W.H., Morton R.A., Epstein J.I., Brooks J.D., Campbell P.A., Bova G.S., Hsieh W.S., Isaacs W.B., and Nelson W.G. 1994. Cytidine methylation of regulatory sequences near the π-class glutathione *S*-transferase gene accompanies human prostatic carcinogenesis. *Proc. Natl. Acad. Sci.* **91:** 11733–11737.

Leu Y.W., Rahmatpanah F., Shi H., Wei S.H., Liu J.C., Yan P.S., and Huang T.H. 2003. Double RNA interference of *DNMT3b* and *DNMT1* enhances DNA demethylation and gene reactivation. *Cancer Res.* **63:** 6110–6115.

Li E., Bestor T.H., and Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915–926.

Liu B., Nicolaides N.C., Markowitz S., Willson J.K., Parsons R.E., Jen J., Papadopolous N., Peltomaki P., de la Chapelle A., Hamilton S.R., et al. 1995. Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat. Genet.* **9:** 48–55.

Lubbert M. 2000. DNA methylation inhibitors in the treatment of leukemias, myelodysplastic syndromes and hemoglobinopathies: Clinical results and possible mechanisms of action. *Curr. Top. Microbiol. Immunol.* **249:** 135–164.

Marks P., Rifkind R.A., Richon V.M., Breslow R., Miller T., and Kelly W.K. 2001. Histone deacetylases and cancer: Causes and therapies. *Nat. Rev. Cancer* **1:** 194–202.

Mizuno K., Osada H., Konishi H., Tatematsu Y., Yatabe Y., Mitsudomi T., Fujii Y., and Takahashi T. 2002. Aberrant hypermethylation of the CHFR prophase checkpoint gene in human lung cancers. *Oncogene* **21:** 2328–2333.

Morin P.J., Sparks A.B., Korinek V., Barker N., Clevers H., Vogelstein B., and Kinzler K.W. 1997. Activation of β-catenin-Tcf signaling in colon cancer by mutations in β-catenin or APC. *Science* **275:** 1787–1790.

Narayan A., Ji W., Zhang X.Y., Marrogi A., Graff J.R., Baylin S.B., and Ehrlich M. 1998. Hypomethylation of pericentromeric DNA in breast adenocarcinomas. *Int. J. Cancer* **77:** 833–838.

Nemoto S., Fergusson M.M., and Finkel T. 2004. Nutrient availability regulates SIRT1 through a forkhead-dependent pathway. *Science* **306:** 2105–2108.

Nguyen C.T., Gonzales F.A., and Jones P.A. 2001. Altered chromatin structure associated with methylation-induced gene silencing in cancer cells: Correlation of accessibility, methylation, MeCP2 binding and acetylation. *Nucleic Acids Res.* **29:** 4598–4606.

Pfeifer G.P., Tang M., and Denissenko M.F. 2000. Mutation hotspots and DNA methylation. *Curr. Top. Microbiol. Immunol.* **249:** 1–19.

Rainier S., Johnson L.A., Dobry C.J., Ping A.J., Grundy P.E., and Feinberg A.P. 1993. Relaxation of imprinted genes in human cancer. *Nature* **362:** 747–749.

Rhee I., Bachman K.E., Park B.H., Jair K.W., Yen R.W., Schuebel K.E., Cui H., Feinberg A.P., Lengauer C., Kinzler K.W., et al. 2002. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* **416:** 552–556.

Rhee I., Jair K.W., Yen R.W., Lengauer C., Herman J.G., Kinzler K.W., Vogelstein B., Baylin S.B., and Schuebel K.E. 2000. CpG methylation is maintained in human cancer cells lacking *DNMT1*. *Nature* **404:** 1003–1007.

Richon V.M. and O'Brien J.P. 2002. Histone deacetylase inhibitors: A new class of potential therapeutic agents for cancer treatment. *Clin. Cancer Res.* **8:** 662–664.

Rideout W.M., III, Coetzee G.A., Olumi A.F., and Jones P.A. 1990. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* **249:** 1288–1290.

Robertson K.D., Ait-Si-Ali S., Yokochi T., Wade P.A., Jones P.L., and Wolffe A.P. 2000. DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat. Genet.* **25:** 338–342.

Romanov S.R., Kozakiewicz B.K., Holst C.R., Stampfer M.R., Haupt L.M., and Tlsty T.D. 2001. Normal human mammary epithelial cells spontaneously escape senescence and acquire genomic changes. *Nature* **409:** 633–637.

Rountree M.R., Bachman K.E., and Baylin S.B. 2000. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat. Genet.* **25:** 269–277.

Sakai T., Toguchida J., Ohtani N., Yandell D.W., Rapaport J.M., and Dryja T.P. 1991. Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am. J. Hum. Genet.* **48:** 880–888.

Sakatani T., Kaneda A., Iacobuzio-Donahue C.A., Carter M.G., de Boom Witzel S., Okano H., Ko M.S., Ohlsson R., Longo D.L., and Feinberg A.P. 2005. Loss of imprinting of *Igf2* alters intestinal maturation and tumorigenesis in mice. *Science* **307:** 1976–1978.

Silverman L.R., Demakos E.P., Peterson B.L., Kornblith A.B., Holland J.C., Odchimar-Reissig R., Stone R.M., Nelson D., Powell B.L., DeCastro C.M., et al. 2002. Randomized controlled trial of azac-

itidine in patients with the myelodysplastic syndrome: A study of the cancer and leukemia group B. *J. Clin. Oncol.* **20:** 2429–2440.

Srinivasan P.R. and Borek E. 1964. Enzymatic alteration of nucleic acid structure. *Science* **145:** 548–553.

Suzuki H., Gabrielson E., Chen W., Anbazhagan R., van Engeland M., Weijenberg M.P., Herman J.G., and Baylin S.B. 2002. A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat. Genet.* **31:** 141–149.

Suzuki H., Watkins D.N., Jair K.W., Schuebel K.E., Markowitz S.D., Chen W.D., Pretlow T.P., Yang B., Akiyama Y., Van Engeland M., et al. 2004. Epigenetic inactivation of *SFRP* genes allows constitutive WNT signaling in colorectal cancer. *Nat. Genet.* **36:** 417–422.

Toyota M., Ho C., Ahuja N., Jair K.W., Li Q., Ohe-Toyota M., Baylin S.B., and Issa J.P. 1999. Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. *Cancer Res.* **59:** 2307–2312.

Ushijima T. 2005. Detection and interpretation of altered methylation patterns in cancer cells. *Nat. Rev. Cancer* **5:** 223–231.

Veigl M.L., Kasturi L., Olechnowicz J., Ma A.H., Lutterbaugh J.D., Periyasamy S., Li G.M., Drummond J., Modrich P.L., Sedwick W.D., and Markowitz S.D. 1998. Biallelic inactivation of *hMLH1* by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proc. Natl. Acad. Sci.* **95:** 8698–8702.

Wales M.M., Biel M.A., el Deiry W., Nelkin B.D., Issa J.P., Cavence W.K., Kuerbitz S.J., and Baylin S.B. 1995. p53 activates expression of HIC-1, a new candidate tumour suppressor gene on 17p13.3. *Nat. Med.* **1:** 570–577.

Wijermans P., Lubbert M., Verhoef G., Bosly A., Ravoet C., Andre M., and Ferrant A. 2000. Low-dose 5-aza-2′-deoxycytidine, a DNA hypomethylating agent, for the treatment of high-risk myelodysplastic syndrome: A multicenter phase II study in elderly patients. *J. Clin. Oncol.* **18:** 956–962.

Wolf P., Hu Y.C., Doffek K., Sidransky D., and Ahrendt S.A. 2001. $O^6$-methylguanine-DNA methyltransferase promoter hypermethylation shifts the *p53* mutational spectrum in non-small cell lung cancer. *Cancer Res.* **61:** 8113–8117.

Wolffe A.P. 2001. Chromatin remodeling: Why it is important in cancer. *Oncogene* **20:** 2988–2990.

Yamashita K., Upadhyay S., Osada M., Hoque M.O., Xiao Y., Mori M., Sato F., Meltzer S.J., and Sidransky D. 2002. Pharmacologic unmasking of epigenetically silenced tumor suppressor genes in esophageal squamous cell carcinoma. *Cancer Cell* **2:** 485–495.

# 附录A 网络资源

## 表观遗传学相关的数据库

| | | |
|---|---|---|
| **染色质和表观遗传学相关网址** | | |
| 染色质结构和功能 | 染色质资源网站，站长为 Jim Bone | www. chromatin. us |
| 杰出表观基因组网络 | 欧洲跨学科表观遗传研究网 | www. epigenome-noe. net |
| HEP | 人类表观基因组计划研究协会 | www. epigenome. org |
| ENCODE | DNA 元件百科：鉴定人类功能元件 | www. genome. gov/12513456 |
| ChromDB | 植物染色质数据库 | www. chromdb. org |
| **组蛋白资源** | | |
| NHGRI 组蛋白数据库 | 含组蛋白序列信息 | research. nhgri. nih. gov/histones |
| Abcam 组蛋白网页 | 含可供下载的组蛋白修饰图谱 | www. abcam. com/chromatin |
| Upstate 组蛋白网页 | 由 Upstate 公司制作的组蛋白修饰图谱 | www. histone. com/modification_map. htm |
| **染色质蛋白** | | |
| NPD：核蛋白数据库 | 收集了已知定位于脊椎动物细胞核中的蛋白 | npd. hgu. mrc. ac. uk/index. html |
| 染色质蛋白 | 与 NPD 网站的搜索引擎相链接，仅限于对染色质蛋白家族成员 | www. epigenome-noe. net/researchtools/structure. php |
| Flybase 染色质网页 | 涵盖果蝇中聚梳、三胸和其它染色质蛋白 | flybase. bio. indiana. edu/allied-data/lk/interactive-fly/aignfam/polycomb. htm |
| **DNA 甲基化** | | |
| 甲基化数据库 | 人类 DNA 甲基化数据库 | www. methdb. de |
| DNA 甲基化 | 专一于生物学甲基化方面的网站 | www. dnamethsoc. com |
| CpG 岛搜索引擎 | CpG 岛序列搜索算法 | www. uscnorris. com/cpgislands2/cpg. aspx |
| **印记** | | |
| 小鼠基因组印记 | 显示已知小鼠印记化区域的染色体图谱 | www. mgu. har. mrc. ac. uk/research/imprinting |
| 基因印记 | 有关基因组印记信息的网站 | www. geneimprint. com/index. html |
| 印记化基因目录 | 印记化基因数据库 | igc. otago. ac. nz/home. html |
| 备选的印记化转录物 | 人和小鼠中预测的印记化转录组资源 | fantom2. gsc. riken. go. jp/imprinting |
| **RNA** | | |
| noncode RNA | 专一于非编码 RNA 的数据库 | www. bioinfo. org . cn/NONCODE |
| RNA 数据库 | 哺乳动物非编码 RNA 数据库 | research. imb. uq. edu. au/rnadb |
| 微 RNA 数据库 | 真核生物微 RNA 数据库 | 166. 111. 30. 65/microrndb |
| 微 RNA 数据库 | 微 RNA 数据库 | microrna. sanger. ac. uk |
| 天然反义 RNA | 天然反义转录物研究平台 | www. narna. ncl. ac. uk |
| ASRP | 拟南芥小 siRNA 和 miRNA 数据资源 | asrp. cgrb. oregonstate. edu |
| **表观遗传资源商业网站** | | |
| Abcam | 组蛋白抗体供应商 | www. abcam. com |
| Upstate | 组蛋白抗体供应商 | www. upstatebiotech. com |
| Epigenomics | DNA 甲基化诊断筛查 | www. epigenomics. com |

## 生物体基因组资源与数据库

| | | |
|---|---|---|
| **所有真核生物** | | |
| NCBI | 多种基因组分析和参考文献的门户网站 | www. ncbi. nlm. nih. gov |
| ENSEMBL | 真核基因组浏览器 | www. ensembl. org |
| UCSC 基因组生物信息学 | 基因组序列及资源门户网站 | genome. ucsc. edu |
| EBI | 各种基因组及蛋白质组计算分析资源门户网站 | www. ebi. ac. uk/services |
| Sanger 研究所 | 序列、生物信息和蛋白质组资源门户网站 | www. sanger. ac. uk/ |
| 贝勒医学院搜索引擎 | 序列搜索及分析工具的门户网站 | searchlauncher. bcm. tmc. edu/docs/sl_links. html |
| zPicture | 序列分析工具 | zpicture. dcode. org |
| iHOP | 蛋白之间信息资源的超级链接 | www. ihop-net. org/UniPub/iHOP |
| Repeat Masker | 重复序列运算法则用以鉴定重复 DNA 序列 | repeatmasker. org |
| Primer 3 | 在线引物设计软件 | frodo. wi. mit. edu/cgi-bin/primer3/primer3. cgi |
| **特定生物** | | |
| 酿酒酵母 | SGD:酵母基因组数据库 ENSEMBL:酿酒酵母基因组分析门户网站 | www. yeastgenome. org www. ensembl. org/Saccharomyces_cerevisiae/index. html |
| 裂殖酵母 | 裂殖酵母相关基因组序列和分析网站门户网站 | www. sanger. ac. uk/Projects/S_pombe |
| 嗜热四膜虫 | TGD:嗜热四膜虫基因组数据库<br>嗜热四膜虫大核基因组测序计划 | www. ciliate. org/<br>www. genome. gov/12512294 |
| 拟南芥* | TAIR:拟南芥信息资源<br>GRAMENE:草类基因组比较<br>水稻基因组资源 | www. arabidopsis. org<br>www. gramene. org/www. tigr. org/tdb/e2kl/osal |
| 秀丽隐杆线虫 | 线虫相关资源门户网站<br>ENSEMBL:线虫基因组分析门户网站线虫相关基因组序列和分析网站门户网站 | www. wormbase. org/www. ensembl. org/Caenorha bd itis_elegans/index. html www. sanger. ac. u k/Projects/C_elega ns |
| 果蝇 | 果蝇基因组数据库<br>ENSEMBL:黑腹果蝇基因组分析门户网站 | www. flybase. net www. ensembl. org/Drosophila_mela nogaster/index. html |
| 爪蟾 | 信息数据库 | www. xenbase. org |
| 小鼠 | 小鼠基因组信息 小鼠品系资源 Baylor 小鼠基因组计划 小鼠 ENSEMBL 资源 | www. informatics. jax. org jaxmice. jax. org/index. html www. mouse-genome. bcm. tmc. edu/ www. ensembl. org/M us_museu I us/index. html |
| 人 | 人类基因组资源门户网站<br>ENSEMBL:人类基因组分析门户网站<br>GDB:人类基因组数据库 | www. ncbi. nlm. nih. gov/genome/guide/human/www. ensembl. org/H om 0 _sa piens/index. html www. gdb. org |

* 应为植物——译者

# 附录 B　组蛋白修饰

## 组蛋白 H2A

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| K5ac | Hs,Sc | Tip60,p300/CBP, Hat1 | 转录激活 | Yamamoto and Horikoshi 1997; Kimura and Horikoshi 1998; Verreault et al. 1998 |
| K9bio | | HCS<br>Biotinidase | 乙酰化与甲基化依赖的<br>参与细胞增殖、基因沉默和细胞对 DNA 损伤的应答 | Stanley et al. 2001;<br>Kothapalli et al. 2005b;<br>Chew et al. 2006 |
| K7ac | Sc | Hat1, Esa1 | 转录激活 | Suka et al. 2001 |
| K13bio | | HCS<br>Biotinidase | 乙酰化与甲基化依赖的<br>参与细胞增殖、基因沉默和细胞对 DNA 损伤的应答 | Stanley et al. 2001;<br>Kothapalli et al. 2005b;<br>Chew et al. 2006 |
| K119ub1 | Dm, Hs | dRing,<br>RING1B | 聚梳沉默<br>紫外损伤应答 | Wang et al. 2004;<br>Kapetanaki et al. 2006 |
| S121ph | Sc | Mec1<br>PIKK | DNA 损伤应答<br>端粒沉默 | Wyatt et al. 2003;<br>Harvey et al. 2005 |
| T125ph | Sc | Mec1<br>PIKK | DNA 损伤应答<br>端粒沉默 | Wyatt et al. 2003; |
| K126bio | Hs | HCS<br>Biotinidase | 乙酰化与甲基化依赖的<br>参与细胞增殖、基因沉默和细胞对 DNA 损伤的应答 | Stanley et al. 2001;<br>Kothapalli et al. 2005b;<br>Chew et al. 2006 |
| K126su | Sc | | 转录抑制<br>阻断组蛋白乙酰化和泛素化 | Nathan et al. 2006 |
| K127bio | Hs | HCS<br>Biotinidase | 乙酰化与甲基化依赖的<br>参与细胞增殖、基因沉默和细胞对 DNA 损伤的应答 | Stanley et al. 2001;<br>Kothapalli et al. 2005b;<br>Chew et al. 2006 |
| S128ph | Sc | Mec1<br>PIKK | DNA 损伤应答<br>端粒沉默 | Downs et al. 2000;<br>Redon et al. 2003;<br>Wyatt et al. 2003;<br>Downs et al. 2004 |
| K130bio | Hs | HCS<br>Biotinidase | 乙酰化与甲基化依赖的<br>参与细胞增殖、基因沉默和细胞对 DNA 损伤的应答 | Stanley et al. 2001;<br>Kothapalli et al. 2005b;<br>Chew et al. 2006 |

注:其他 H2A 修饰: S1ph, K4ac, K13me, K15ac, K21ac, K36ac, K74me, K75me, R77me, K95me, T120ph, H2A 崇基化(Pantazis and Bonner 1981; Song et al. 2003; Alhara et al. 2004)

## 组蛋白 H2B

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| K5ac | Hs | | 转录激活 | Puerta et al. 1995; Galasinski et al. 2002 |
| S10ph | Sc | Ste20 | 凋亡 | Ahn et al. 2005 |
| S14ph | Hs, Mm | Mst1/krs2 | 凋亡 | Ajiro 2000; Cheung et al. 2003 |
| | | 激酶 | 体细胞高突变和类别转换重组 | Odegard et al. 2005 |
| K16su | Sc | | 基因抑制 | Nathan et al. 2006 |
| K17su | Sc | | 基因抑制 | Nathan et al. 2006 |
| S33ph | Dn | CTK | 转录激活 | Maile et al. 2004 |
| K120ub | Hs | | 通过 H3 甲基化来与激活转录的 SAGA 协作以调控细胞周期进程,减数分裂 | Robzyk et al. 2000; Sun and Allis 2002; Kao et al. 2004 |
| K123ub | Sc | Rad6(E2) | 通过降低组蛋白 H3K4 和 H3K79 位的甲基化导致端粒沉默 | Emre et al. 2005 |
| | | Brel(E3)-ubl | | |

注:其他的 H2B 修饰:E2arn,K5me,K6su,K7su,K11ac,K12ac,K15ac,K16ac,K20ac,K23me,K24ac,S32ph,K43me,K85ac,R99me,K108ac,K116ac,H2B 糖基化,H2B 生物素化(Rouleau et al. 2004)

## 组蛋白 H3

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| R2me | Hs | CARM1-me2a | 基因表达 | Chen et al. 1999; Schurter et al. 2001 |
| | Mm | | | |
| T3ph | Hs | Haspin | 着丝粒纺锤体功能 | Polioudaki et al. 2004; Dai et al. 2005 |
| | At | | | |
| K4me | Sc | Set1-me3 | rDNA 沉默,端粒沉默,转录激活 | Briggs et al. 2001; Roguev et al. 2001; Na et al. 2002; Bryk et al. 2002; Bernstein et al. 2002; Santos-Rosa et al. 2002 |
| | Tt | | 转录激活 | Strahl et al. 1999 |
| | Hs | SET7/Set9-mel | 转录激活 | H. Wang et al. 2001 a; Nishioka et al. 2002a; Wilson et al. 2002; Zegerman et al. 2002 |
| | | MLL | | |
| | Ds | MLL2-me3 | 三胸活化 | Milne et al. 2002; Nakamura et al. 2002 |
| | Hs | MLL3 | | |
| | | Ash1-me2 | | |
| | Ds | | 三胸活化 | Beisel et al. 2002; sanchez-Elsner et al. 2006 |
| | Hs | SMYD3-me3 | | |
| | Hs | Meisetz-me3 | 转录激活 | Hamamoto et al. 2004 |
| | Mm | | 有丝分裂早期进程 | Hayashi et al. 2005 |
| K9ac | Sc | SAGA | 转录激活 | Grant et al. 1999 |
| | Hs | SRC1 | 核受体共激活因子 | Spencer et al. 1997; Schubeler et al. 2000; Vaquero et al. 2004 |
| | Dm | | 转录激活 | Nowak et al. 2000 |

续表

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| K9me | Mm, Hs | G9a-me1,me2 | 转录抑制<br>印记 | Tachibana et al. 2001, 2002; Ogawa et al. 2002; Xin et al. 2003 |
| | Dm | Su(var)3-9-me2 | 显性位置花斑效应修饰 | Czermin et al. 2001; Schotta et al. 2002; Ebert et al. 2004 |
| | Mm | Suv39h 1-me3<br>Suv39h2-me3 | 着丝粒周边异染色质 | OCarroll et al. 2000; Rea et al. 2000; Lachner et al. 2001; Peters et al. 2001 |
| | Hs | SUV39H1-me3 | Rb 介导的基因沉默 | Nielsen et al. 2001; Vandel et al. 2001 |
| | Sp | Clr4-me1,me2 | 着丝粒和接合型沉默 | Bannister et al. 2001; Nakayama et al. 2001 |
| | Nc | Dim5-me3 | DNA 甲基化 | |
| | At | KRYPTONITE-me2 | DNA 甲基化 | Tamaru and Selker 2001 |
| | Hs | EuHMTase1-me1,me2 | 转录抑制 | Jackson et al. 2002, 2004 |
| | Hs, Mm | ESET-me2,me3 | 转录抑制 | Ogawa et al. 2002; Tachibana et al. 2005 |
| | Dm, Hs | Ash1-me2 | 三胸激活 | Schultz et al. 2002; Yang et al. 2002; |
| | Hs | RIZ1-me2 | 肿瘤抑制,雌激素应答 | Dodge et al. 2004; Wang et al. 2004 Beisel et al. 2002 Kim et al. 2003; Carling et al. 2004 |
| S10ph | Sc | Snf1 | 转录激活 | Lo et al. 2001 |
| | Dm | Jil-1 | 雄性 X 染色体转录上调 | Jin et al. 1999; Y. Wang et al. 2001 |
| | Hs | Rsk2<br>Msk1<br>Msk2 | 早期基因的转录激活（与 H3K14 乙酰化相协作） | Sassone-Corsi et al. 1999; Thomson et al. 1999; Cheung et al. 2000; Clayton et al. 2000 |
| | Hs | Ikkα | 转录上调 | Anest et al. 2003; Yamamoto et al. 2003 |
| | Sc, Ce | Ipl1/AuroraB | 有丝分裂染色体致密化 | Hendzel et al. 1997; Wei et al. 1999; Hsu et al. 2000 |
| | An | NIMA | 有丝分裂染色体致密化 | De Souza et al. 2000 |
| | Hs, Ce | Fyn kinase | UVB 诱导的 MAP 激酶通路 | He et al. 2005 |
| T11ph | Hs | Dlk/ZIP | 有丝分裂特异性磷酸化 | Preuss et al. 2003 |
| K14ac | Sc,Tt,Mm | Gcn5 | 转录激活 | Brownell et al. 1996; Kuo et al. 1996 |
| | Hs, Om | $TAF_{II}230$<br>$TAF_{II}250$ | 转录激活 | Mizzen et al. 1996 |
| | Hs | p300 | 转录激活 | Schiltz et al. 1999 |
| | Hs | PCAF | 转录激活 | Schiltz et al. 1999 |
| | Mm | SRC1 | 核受体共激活因子 | Spencer et al. 1997 |
| R17me | Hs, Mm | CARM1 | 转录激活（与 H3K18/23 乙酰化协作） | Chen et al. 1999; Schurter et al. 2001; Bauer et al. 2002; Daujat et al. 2002 |
| K18ac | Sc | SAGA<br>Ada | 转录激活 | Grant et al. 1999 |
| | Hs | p300 | 转录激活 | Schiltz et al. 1999 |
| | Hs | CBP | 转录激活（与 H3R17 甲基化协作） | Daujat et al. 2002 |

续表

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| K23ac | Sc | SAGA | 转录激活 | Grant et al. 1999 |
| | Hs | CBP | 转录激活(与 H3R17 甲基化协作) | Daujat et al. 2002 |
| R26me | Hs | CARM1 | 体外甲基化位点 | Chen et al. 1999; Schurter et al. 2001 |
| K27me | Hs, Dm | E(z)/ EZH2-me3 | 聚梳抑制<br>早期 B 细胞发育<br>X 染色体失活 | Cao et al. 2002; Czermin et al. 2002; Kuzmichev et al. 2002; Muller et al. 2002;<br>Su et al. 2003 |
| S28ph | Hs | Aurora-B | 有丝分裂染色体致密化 | Goto et al. 1999,2002 |
| | Hs | MSK1 | UVB 诱导的磷酸化 | Zhong et al. 2001 |
| K36me | Sc | Set2-me2 | 基因抑制 | Strahl et al. 2002; Kizer et al. 2005 |
| | Nc | Set2-me2 | 转录激活 | Adhvaryu et al. 2005 |
| | Sp | Set2-me2 | 转录延伸 | Morris et al. 2005 |
| K79me | Sc,Hs | Dot1/DOT1 L-me2 | 端粒沉默,粗线期检验点 | Feng et al. 2002; Lacoste et al. 2002; Ng et al. 2002; van Leeuwen et al. 2002 |

注:其他组蛋白 H3 修饰:K14me, K23me, K27ac, T32ph, K37me, K56me, K64me, K115ac. K118ac. K118me, K122ac, R128me (Hyland et al. 200S).

## 组蛋白 H4

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| S1ph | Hs, Sc | 酪蛋白激酶 II | DNA 损伤应答 | Ruiz-Carrillo et al. 1975; Cheung et al. 2005;<br>van Attikum and Gasser 2005 |
| R3me | Hs, Sc | PRMT1 | 转录激活 | H. Wang et al. 2001 a,b |
| K5ac | Tt, Dm, Hs | Hat1 | 组蛋白装配 | Sobel et al. 1995; Parthun et al. 1996; Tap lick et al. 1998; Turner 2000; Kruhlak et al. 2001. |
| | Sc | Esa1/NuA4 | 细胞周期进程 | Smith et al. 1998; Allard et al. 1999; Clarke al. 1999; Bird et al. 2002; Miranda et al. 2006 |
| | Hs, Mm | ATF2 | 序列特异性转录因子 | Kawasaki et al. 2000a |
| | Hs | p300 | 转录激活 | Turner and Fellows 1989; Schiltz et al. 1999 |

续表

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| K8ac | HS,Mm | Y - ATF2 | Xi 排斥<br>序列特异性转录因子 | Jeppesen et al. 1993; Choy et al. 2001; Kruhlak et al. 2001; Kawasaki et al. 2000b |
| | Hs | PCAF/p300 | 转录激活 | Turner and Fellows 1989; Schiltz et al. 1999 |
| K12ac | Sc, Hs | Hat1 | Xi 排斥<br>组蛋白组装 | Turner and Fellows 1989; Jeppesen et al. 1993; Kleff et al. 1995; Sobel et al. 1995; Parthun et al. 1996; Chang et al. 1997; Kruhlak et al. 2001 |
| | Sc | NuA4 | 有丝分裂和减数分裂进程 | Choy et al. 2001 |
| K12bio | Hs | HCS<br>Biotinidase | DNA 双链断裂后此修饰减少<br>细胞增殖 | Stanley et al. 2001; Kothapalli et al. 2005a,b |
| K16ac | Mm | | Xi 排斥<br>细胞周期依赖的乙酰化<br>雄性 X 染色体的转录上调 | Jeppesen et al. 1993; Taplick et al. 1998<br>Akhtar and Becker 2000; Hsu et al. 2000 |
| | Dm | MOF | 序列特异性转录因子 | Turner and Fellows 1989; Kawasaki et al. 2000a; Turner 2000; Kruhlak et al. 2001; Vacquero et al. 2004 |
| | Hs, Mm | ATF2 | | |
| K20me | Mm,Dm | Suv4-20h1-me2,me3<br>Suv4-20h2-me2,me3 | 基因沉默 | Schotta et al. 2004 |
| | Hs, Dm | Pr-SET7/Set8-me1 | 转录沉默,有丝分裂致密化 | Fang et al. 2002; Nishioka et al. 2002b; Rice et al. 2002<br>Beisel et al. 2002 |
| | Dm | Ash1-me2 | 与 H3K4 和 H3K9 甲基化协作的三胸激活 | |
| K59me | Sc | | 沉默染色质的形成 | Zhang et al. 2003 |
| Su | Hs | SUMO-1<br>SUMO-3 | 转录抑制 | Shiio and Eisenman 2003 |

注:其他组蛋白 H4 修饰:K12me, S47ph, K31ub1, K59me, K77ac, K79ac, K79me, K91ub1, R92me (Hyland et al. 2005).

## 组蛋白 H1

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| E2ar1 | Rn | PARP-1 | 神经营养活性相关 | Ogata et al. 1980; Visochek et al. 2005 |
| T10ph | Hs | | 有丝分裂特异性<br>H1b 的转录激活 | Chadee et al. 1995; Garcia et al. 2004; Sarg et al. 2006 |
| E14ar1 | Rn | PARP-1 | 神经营养因子活性相关 | Ogata et al. 1980; Visochek et al. 2005 |

续表

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| S17ph | Hs | | 细胞间期特异性<br>H1b 的转录激活 | Garcia et al. 2004; Chadee et al. 1995; Sarg et al. 2006 |
| K26me | Hs | EZH2-me2 | 介导 HP1 的结合 | Kuzmichev et al. 2004; Daujat et al. 2005 |
| S27ph | Hs | | 阻断 HP1 的结合 | Garcia et al. 2004; Daujat et al. 2005 |
| T137ph | Hs | | 有丝分裂特异性<br>H1b 的转录激活 | Chadee et al. 1995; Garcia et al. 2004; Sarg et al. 2006 |
| T154ph | Hs | | 有丝分裂特异性<br>H1b 的转录激活 | Chadee et al. 1995; Garcia et al. 2004; Sarg et al. 2006 |
| S172ph | Hs | | 细胞间期特异性<br>H1b 的转录激活 | Chadee et al. 1995; Garcia et al. 2004;Sarg et al. 2006 |
| S188ph | Hs | | 细胞间期特异性<br>H1b 的转录激活 | Chadee et al. 1995; Garcia et al. 2004; Sarg et al. 2006 |
| K213arl | Rn | PARP-1 | 神经营养因子活性相关 | Ogata et al. 1980; Visochek et al. 2005 |

注:其他组蛋白 H1 修饰:K26ac, E114arn, H1 ubiquitination, H1 carbonylation (Pham and Sauer 2000; Rouleau et al. 2004).

## 组蛋白 H2AX

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| S139ph | Hs, Sc<br>Dm,Xl | ATM<br>DNA-PK | DNA 修复<br>M 期相关<br>亦被称为 γH2AX | Rogakou et al. 1998,1999; Burma et al. 2001; Stiff et al. 2004; Ichijima et al. 2005; Mukherjee et al. 2006 |

注:其他 H2AX 修饰:T136ph

## 组蛋白 macroH2A

| 修饰位点 | 修饰模式 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| K17me | Hs | | | Chu et al. 2006 |
| K115ubl | Hs | | | Ogawa et al. 2005; Chu et al. 2006 |
| K122me | Hs | | | Chu et al. 2006 |
| T128ph | Hs | | | Chu et al. 2006 |
| K238me | Hs | | | Chu et al. 2006 |

## 组蛋白 H3.3

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| K4me | Dm | mel,me2,me3 | 转录激活 | Mckittrick et al. 2004 |
| K9me | Dm | mel,me2, | 转录抑制 | Mckittrick et al. 2004 |
| K9ac | Dm ,Hs | | 转录激活 | Mckittrick et al. 2004; Hake and Allis. 2006 |
| K14me | Dm | mel,me2, | | Mckittrick et al. 2004 |
| K14ac | Dm ,Hs | | 转录激活 | Mckittrick et al. 2004; Hake and Allis. 2006 |
| K18ac | Hs | | 转录激活 | Hake and Allis. 2006 |
| K23AC | Hs | | 转录激活 | Hake and Allis. 2006 |
| K27me | Dm | mel,me2,me3 | 转录抑制 | Mckittrick et al. 2004 |
| K36me | Dm, Hs | mel,me2,me3 | 转录激活 | Mckittrick et al. 2004; Hake and Allis. 2006 |
| K37me | Dm | mel,me2, | | Mckittrick et al. 2004 |
| K79me | Dm, Hs | mel,me2, | 转录激活 | Mckittrick et al. 2004; Hake and Allis. 2006 |
| S31ph | Mammals | | 有丝分裂特异性磷酸化 | Hake et al. 2005 |

## CEN-H3/CENP-A

| 修饰位点 | 模式生物 | 修饰酶 | 功能 | 参考文献 |
|---|---|---|---|---|
| S7ph | Hs | | 有丝分裂 | Zeitlin et al. 2001 |

注:模式生物名称缩写:

(An)*aspergillus nidulan*,构巢曲霉;(At)*Arabidopsis thaliana*,拟南芥;(Ce)*Caenorhabditis elegan*,秀丽隐杆线虫;(Dm)*Drosophila melanogaster*,黑腹果蝇;(Hs)*Home sapiens*,人;(Mm)*Mus musculus*,小鼠;(Nc)*Neurospora crassa*,链孢霉;(Rn)*Rattus norvegicus*,褐家鼠;(Sc)*Saccharomyces cerevisiae*,酿酒酵母;(Sp)*Schizosaccharomyces pombe*,裂殖酵母;(Tt)*Tetrahymena thermophila*,嗜热四膜虫;(Xl)*Xenopus laevis*,非洲爪蟾。

1. 组蛋白修饰依据 Turner 提议的命名法进行命名。

2. 图中列出了所有已知的组蛋白修饰、修饰酶以及它们的原始参考文献(截至 2006 年 5 月)。其他未知功能的修饰罗列在图表下方。这些修饰信息通过多种渠道获得,包括综述文章,来自 Abcam cambridge、UK 和 Upstate Charlottesville、USA 的信息以及 Reinberg 实验室未发表的一些数据。

3. 本表是在 Lachner 等于 2003 年建立的图表的基础之上编纂而成的,Roopsha Sengupta 和 Mario Richter(维也纳分子病理学研究院 Jenuwein 实验室)丰富了图表内容。Patrick Trojer 博士(新泽西休斯医学研究院 Reinberg 实验室)进行了核实。

4. Steven Gray 博士(柏林 St. James 医院药物分子研究院临床药物部)提供的信息对本表进行了验证及补充。

张 慧 译
吴 慧 校
朱 冰 审校

# 附录C　不同模式生物中的表观遗传机制

| | | 酿酒酵母 | 裂殖酵母 | 链孢霉 | 秀丽隐杆线虫 | 果蝇 | 哺乳动物 | 拟南芥 |
|---|---|---|---|---|---|---|---|---|
| **基因组特征** | | | | | | | | |
| | 基因组大小 | 12Mb | 14Mb | 40Mb | 100Mb | 180Mb | 3 400Mb | 150Mb |
| | 基因数目 | 6 000 | 5 000 | 10 000 | 20 000 | 14 000 | 约 25 000 | 25 000 |
| | 基因平均大小 | 1.45kb | 1.45kb | 1.7kb | 2kb | 5kb | 35～46kb | 2kb |
| | 内含子/基因的平均数 | ≤1 | 2 | 2 | 5 | 3 | 6～8 | 4～5 |
| | 编码蛋白的基因组比例 | 70 | 60 | 44 | 25 | 13 | 1～1.5(*Hs*) | 26 |
| **表观遗传特征** | | | | | | | | |
| ON | 组蛋白乙酰化 | + | + | + | + | + | + | + |
| ON | H3K4 甲基化 | + | + | + | + | + | + | + |
| ON | H3K36 甲基化 | + | + | + | + | + | + | + |
| ON | H3K79 甲基化 | + | + | + | + | + | + | + |
| ON | H3.3 组蛋白变体 | + | + | + | + | + | + | + |
| ON/OFF | SWI/SNF ATP 酶复合物 | + | + | + | + | + | + | + |
| ON | CHD1 ATP 酶家族 | + | (+)[a] | (+)[a] | (+)[a] | + | + | + |
| ON | SWR1 ATP 酶 | + | (+)[a] | (+)[a] | (+)[a] | + | + | (+)[a] |
| ON/OFF | ISWI ATP 酶 | + | + | + | + | + | + | + |
| ON/OFF | INO80 ATP 酶 | + | + | + | + | + | + | + |
| OFF | MI-2 ATP 酶 | − | (+)[a] | (+)[a] | + | + | + | + |
| OFF | CENP-A 着丝粒组蛋白变体 | + | + | + | + | + | + | + |
| OFF | H3K9 甲基化[b] | − | + | + | + | + | + | + |
| OFF | HP1 类似蛋白 | − | + | + | + | + | + | + |
| OFF | RNA 干扰 | − | + | + | + | + | + | + |
| OFF | H4K20 甲基化[c] | − | + | + | + | + | + | + |
| OFF | H3K27 甲基化 | − | − | + | + | + | + | + |
| OFF | polycomb 抑制复合物 | − | − | − | + | + | + | + |
| OFF | DNA 甲基化 | − | − | + | − | (+)[d] | + | + |
| OFF | DNA 甲基化结合蛋白 | − | +[e] | + | +[f] | +[g] | + | + |
| OFF | 基因印记 | − | − | − | − | +[h] | + | + |

注：*Hs*，*Homo sapiens*；

a 依据其序列同源性认为存在但是功能未知的表观遗传特征。

b 有证据显示 H3K9 甲基化存在于活跃染色质区域，然而功能重要性未知。

c 酿酒酵母中不存在 H4K20 的三甲基化，然而 H4K20 的三甲基化状态存在于多细胞生物体中。

d 果蝇中的 DNA 甲基化水平非常低。

e 突变的 Dnmt2

f Dnmt2(Pp)和 MBD 结构域蛋白(Ce，Cb，Pp)

g Dnmt2 和 MBD 结构域蛋白(Dm)

h 染色质和基因组范围内的，而非基因特异性的

# 附录 D　组蛋白修饰的 *Brno* 命名法

| 修饰基团 | 被修饰氨基酸 | 修饰水平 | 修饰的缩写[a] | 修饰残基举例[b] |
|---|---|---|---|---|
| 乙酰基 | 赖氨酸 | 单 | ac | H3K9ac |
| ADP 核糖 | 谷氨酸 | 单 | arl | H2BE2arl |
| | 谷氨酸 | 多 | arn | H2BE2arn[d] |
| 生物素 | 赖氨酸 | 单 | bio | H4K12bio |
| 甲基 | 精氨酸 | 单甲基化 | mel | H3R17mel |
| | 精氨酸 | 对称二甲基化 | me2s | H3R2me2s |
| | 精氨酸 | 不对称二甲基化 | me2a | H3R17me2a |
| 甲基 | 赖氨酸 | 单 | mel | H3K4mel |
| | 赖氨酸 | 二 | me2 | H3K4me2 |
| | 赖氨酸 | 三 | me3 | H3K4me3 |
| 磷酸基团 | 丝氨酸或苏氨酸 | 单 | ph | H3S10ph |
| 泛素 | 赖氨酸 | 单[c] | Ubl | H2BK120ubl |
| Sumo 基团 | 赖氨酸 | 单 | su | H4K5su[d] |

a 修饰采用小写字母，以便将其与大写的氨基酸（单个字母编码的）或组蛋白（如 H2A）区分开来。

b 命名原则是从左起依次为组蛋白、残基、修饰。一旦被修饰残基未知或不相关，在组蛋白后加上修饰即可，例如 H4ac 和 H2Bbarl。多重修饰只需根据需要简单的罗列出来（如 H3K4me3K9acS10ph）。因为每个修饰残基均以大写字母开头来界定特定氨基酸，并且修饰被指定为小写字母，因此在多重修饰的情况下，无需用逗号或圆点来分隔不同修饰的残基。有时，没有修饰的残基有可能是残基信息承载联合体的重要组成部分，在这种情况下这个残基应该被添加上（如 H3K9S10ph）以表明 H3 第 9 位赖氨酸残基未被修饰，第 10 位丝氨酸残基被磷酸基团修饰。

c 多泛素化组蛋白被缩写为 ubn

d 目前假定存在的

获许并改编自 Turner B. M.，Nat. Struct. Mol. Biol. 12：110～112（2005）

# 索　引


(TCR) β　474
5-氮杂胞苷　413
AID　481
*ALD*　515
AML 综合征　543
Angelman 综合征　518
Argonaute　130
ATP 依赖的染色体重塑体　39
BCR　474
Beckwith-wiedemann 综合征　518
BECKWITH-WIEDEMANN 综合征　521
Blimp1　482
Brahma　48
*Cd79a*　473
CenH　128
CENP-A　124，127，134
Chp1　127
Chp2　127
chromo 结构域　127
Cnp1　127
CpG 岛　49
*c-Kit*　471
Dicer　130
DNA 低甲基化　62
DNA 高甲基化　63
DNA 甲基化　122，447
DNA 剔除　52
Dnmt3a　553
Dnmt3b　553
*DNMT*　554
DPY　349
E2A　474
EBF1　474
Esc1　90
ESC　358
Ezh2　476
FACT　48
H2A. Z　88
H3K27 甲基化　476
H3K9　123
H3K9 二甲基化　552
H4K16　87，90
HDAC　129
*her-1*　348
*HML*　83
*HMLα*　79
*HMR*　83
*HMR*a　79
HP1　40
*IFNBI*　483
*IL-7* 受体　471
LSD1　45
*MAT*　79
mb-1　474
MES 蛋白　357
MIX-1　349
MLL　57
MSUD　136，143
Muller 棘轮　381
*Notch1* 基因　472
Pax5　471
PcG　56
PML　57
PML 综合征　543
Polycomb 应答元件　56
Prader-Willi 综合征　518
PRC2　358
*P* 元件　100
Rap1　87，88，90
RDRC　130
Rett 综合征　524
RIP　136
RITS　130
RNAi　129，130
RNA 干扰　122，123
RNA 干涉机制　455
RNA 聚合酶　501
RNA 聚合酶 IV　53

RNA 依赖的 RNA 聚合酶 130
RNA 诱导沉默复合体 114
Rubinstein-Taybi 综合征 523
Sas2 87
Schimke immuno-osseous 发育异常综合征 527
SET 结构域蛋白 141
shmoo 78
Silver-Russell 综合征 518
silver-russell 综合征 522
SIR1 81
SIR2 81
SIR3 81
SIR4 81
SIR 90
Sir 沉默复合体 83
SIR 复合体 87
SUMO 242
Swi6 127
TAS 重复 129
trxG 56
TSA 134
V（D）J 重排 474
XOL-1 348
X 染色体失活（XCI） 525
X 染色体失活 58，504
X 染色体失活和印记 501
X 染色体失活中心 58，385
X 信号元件（XSE） 350
yKu70/yKu80 85
yKu 88，91
“Knudson 双重打击”理论 62
α 地中海贫血症 X 连锁的智力迟钝 526
癌基因（oncogene） 62
癌症 500
巴氏小体 383
半保留（复制）过程 66
半甲基化 50
保守 452
表观等位基因 32
表观生成 449
表观突变 463
表观遗传 447，491，497
表观遗传基因型 515
表观遗传记忆 501
表观遗传密码 35
表观遗传学全貌 35
表观遗传重编程 504
草履虫 31
常染色质 37
成纤维细胞 498
池内 A 颗粒 457
雌雄同体 347
次甲基四氢叶酸还原酶缺陷 528
从头甲基化 50
脆性 X 染色体综合征 529
达尔文进化理论 29
单亲双体型 425
单亲源二倍体 517
蛋白 349
灯刷状染色体 501
等位基因排斥 479
点突变 62
动粒染色质 124
动态突变 516
端粒 504
端粒位置效应 77，105
端粒相关卫星序列 105
端粒相关序列 113，128
多莉 60
多能胚胎干细胞（ES） 498
多能性 447
多能性源祖细胞 469
多能状态 499
儿童早老症 58
繁殖性克隆 506
反式效应 38
反义转录 477
泛素 242
防卫机制 44
纺锤丝 494
非编码 RNA 130
非合子子囊 123
非配对 DNA 导致的减数分裂沉默 135
非配对 DNA 的减数分裂沉默（MSUD） 354
非整倍化 502
非组蛋白类染色质蛋白 447
父本基因组 458
改版核移植 508
干扰素 450
睾丸支持细胞 498

骨髓细胞 471
骨髓增生异常综合征 555
合子 501
合子发育控制 493
核二形性 31
核小体重塑 67
核移植 491，498
核移植胚胎干细胞 494
核移植治疗 506
核重编程 501
核转移 62
活化诱导的胞嘧啶核苷脱氨酶 481
活性染色质区域 44
基因印记缺失 548
基因印记中心（IC） 520
基因印记中心 519
基因转录水平沉默 130
基因组 492
基因组的全能性 55
基因组印记 492
基因座 454
剂量补偿 311
剂量补偿复合物 60，348
剂量补偿效应 348，504
甲基化 CpG 结合结构域蛋白 50
甲基化模式 503
甲基磺酸乙酯 103
假孕 494
兼性异染色质 37，58
减数分裂 455，492
减数分裂后期 494
减数分裂前诱导的甲基化 139
减数分裂驱动 336
碱性磷酸酶 450
浆细胞 451
接合型转换 31
姐妹综合征 518
精氨酸甲基化 45
精蛋白 457
精子 501
抗原受体基因 474
可变表面抗原基因 5
克隆动物 491
克隆技术 506
赖氨酸特异性“去甲基化酶” 45
裂殖酵母 52
磷酸酶 45
卵黄囊 458
卵裂 492，500
卵母细胞 455，492，494
卵丘细胞 498
螺线管 40
免疫排斥 506，507
免疫球蛋白 477
免疫缺陷 527
免疫系统 447
免疫抑制药物 507
面部异常综合征 527
面肩胛臂营养失调 530
模式系统 31
母本基因组 458
母本效应 168
母源发育控制 493
母源决定性因子 456
囊胚 448，455，491，498
内胚层 448，491
内胚层症状 505
内细胞团 448
拟常染色体区域 384
黏接蛋白 92
黏结蛋白 132
尿囊 450
胚泡 494，496，501
胚胎着床 448
胚外组织 448
配子 503
配子差异性 DNA 甲基化区域 433
配子甲基化印记 433
平衡罗伯逊易位 518
平抑 135，136
起始位点识别复合体 81
嵌合体 460，494
亲本 492
亲本基因组 503
曲古抑菌素 A 130
去甲基化 503
去收缩 479
去乙酰化酶 211
全能性 447，492
全能状态 499

染色体异常 496
染色质 501
染色质附属因子 413
染色质免疫沉淀实验 46
染色质免疫共沉淀 88
染色质修饰 497
染色质重塑 481
染色质重塑复合物 47，502
胚粒 29
散漫着丝粒染色体 325
桑葚胚 448
桑椹胚 492
神经外胚层 505
生殖系 497
生殖系干细胞的自我更新 455
生殖细胞 449
生殖细胞前体 448
生殖原细胞 449
生殖原细胞世系 450
世系化决定 471
受精过程 448
受精卵 491，492
受体 500
受体选择 500
双元开关 41
顺式效应 38
宿主 447
宿主针对移植物 506
酸性结构域 309
随机失活 504
髓系共同源祖细胞 471
胎盘 448，503，504
体细胞高突变 51
体细胞核移植 494
同源异型 33
同宗配合 79
透明带 494
突变 447
外胚层 448
完整性 494
微球菌核酸酶 101
微阵列芯片 547
伪常染色体区 381
位点收缩 477
位置效应花斑 2，33，81，99
细胞膜 494
细胞世系决定 448
细胞质内精子注射 520
细胞周期 48，498，499
锌指结构 452
新着丝粒 41
修饰盒 46
秀丽隐杆线虫 346
溴区结构域 39
嗅神经元细胞 499
亚端粒元件 Y′ 89
亚端粒重复元件 X 89
亚核重定位 477
视网膜母细胞瘤基因 62
遗传标记 494
遗传潜能 495
遗传信息编码 447
遗传重排 498
遗传转变 500
乙酰化 475
乙酰基转移酶（HAT） 87
乙酰转移酶 211
异染色质 37
异染色质蛋白 1 122
异染色质空间 112
有丝分裂 455，502
预世系化 472
预形成 449
原始生殖细胞 448
杂合性丢失 547
治疗性核移植 491
治疗性克隆 494
致密因子 349
致密因子复合物 234
中胚层 448
重编程 447，491，496，498，499，501
重编程能力 498
重复序列 503
重复序列诱导的点突变 6，31，135
重复元件 54
转录沉默 453
转录后基因沉默 131
转录后水平基因沉默 52
转录水平 501
转录水平基因沉默 52

转录因子 42
着丝粒区域不稳定 527
滋养外胚层 448，503
子宫 497
自然选择 29
阻遏蛋白激活蛋白 81
组成型异染色质 37，58
组蛋白 H3K9 甲基化 122
组蛋白变体 47，48，447
组蛋白甲基化状态 503
组蛋白甲基转移酶 107
组蛋白精氨酸甲基转移酶 45
组蛋白赖氨酸甲基转移酶（HKMT） 451
组蛋白密码 476
组蛋白去乙酰化酶 127，543，552
组蛋白修饰 447，448

[General Information]
书名=表观遗传学
作者=（美）Allis等编著
页数=576
出版社=科学出版社
出版日期=2009.7
SS号=12358215
DX号=000006762884
url=http://book2.duxiu.com/bookDetail.jsp?dxNumber=000006762884&d=FC78EA011D802694C459D1E556AD789D&fenlei=15080605&sw=%B1%ED%B9%DB%D2%C5%B4%AB%D1%A7

封面
书名
版权
前言
目录
第1章　表观遗传学：从现象到领域
第2章　表观遗传学发展简史
第3章　概述与概念
第4章　酿酒酵母的表观遗传学
第5章　果蝇位置效应花斑、异染色质形成和基因沉默
第6章　表观遗传学研究的真菌模型：裂殖酵母和链孢霉
第7章　纤毛动物的表观遗传学
第8章　RNAi和异染色质组装
第9章　植物的表观遗传调控
第10章　染色质修饰及其作用机理
第11章　Polycomb蛋白家族调控的转录沉默
第12章　三胸蛋白家族（trxG）与转录调控
第13章　组蛋白变体和表观遗传学
第14章　染色体遗传的表观遗传学调控
第15章　秀丽隐杆线虫X染色体的表观遗传学调控
第16章　果蝇中的剂量补偿效应
第17章　哺乳动物中的剂量补偿效应
第18章　哺乳动物的DNA甲基化
第19章　哺乳动物基因组印记
第20章　生殖细胞系和多能干细胞
第21章　淋巴细胞生成的表观遗传学调控
第22章　核移植和基因组重编程
第23章　表观遗传学和人类疾病
第24章　癌症的表观遗传决定因素
附录A网络资源
附录B组蛋白修饰
附录C不同模式生物中的表观遗传机制
附录D组蛋白修饰的Brno命名法
索引